Development of Approaches for Investigating the Distribution of *Toxoplasma gondii* Infection in Natural Populations of Animals and Humans

Jaroslav Bajnok

School of Environment and Life Sciences

University of Salford, Salford, UK

Submitted in Fulfilment of the Requirements of the Degree of Doctor of Philosophy

January 2017

TABLE OF CONTENTS

List of Figures ……………………………………………………………………………IV

List of Tables ……………………………………………………………………..…VI

Acknowledgments…………………………………………………………………VIII

Abbreviations ………………………………………………………………….....IX

Declaration ………………………………………………………………………XI

Abstract ……………………………………………………………………………XII

Chapter 1: General Introduction 1

1. Introduction…………………………………………………………….………..2

1.1. Life cycle and morphology …………………………………………….…….…3

1.2. Transmission …………………………………………………….………....…8

1.3. Human toxoplasmosis……………………………………………………….…10

1.3.1. Diagnostics …………………………………………………………......12

1.4. The ecology of the wood mouse (*Apodemus sylvaticus*) ……………………......16

1.5. Malham Tarn …………………………………………………………….…17

1.6. Lung cancer ………………………………………………………….......21

1.7. Chronic obstructive pulmonary disease (COPD) ……………………………......22

Chapter 2: Materials and Methods ……………………………………………….25

2.1. Sampling of *Apodemus sylvaticus*…………………………………………..…26

2.2. PCR detection of *Toxoplasma gondii* and RFLP genotyping ....…...27

2.2.1 Mammalian α-tubulin PCR …………………………………………..…28

2.2.2 Amplification of the B1 Gene ……………………………………..……28

2.2.3 SAG1 PCR ……………………………………………………………...29

2.2.4 SAG2 PCR …………………………………………………………....30

2.2.5 SAG3 PCR …………………………………………………………....31

2.2.6 GRA6 PCR …………………………………………………………...32

2.3. Microsatellites analysis and host genotyping ……………………………….....33

2.4. Processing and analysis of human lung tissue samples ……………………….…39

2.4.1. Haematoxylin and Eosin staining …………………………………………………..39

2.4.2 Immunohistochemistry (IHC) for *T. gondii* in human lung sections 40

2.5 PCR detection and genotypes of *T. gondii* in human samples ……………………....41

Chapter 3: Microsatellite-based genetic diversity and population structure of *Apodemus sylvaticus* in Malham Tarn Nature Reserve ……………………………………………....43

3.1. Introduction……………………………………………………………………….....44

3.2. Materials and Methods ...…………………………………………………………....47

3.3. Results …………………………………………………………………………….....48

3.4. Discussion ………………………………………………………………………….…60

Chapter 4: An investigation into the prevalence of *Toxoplasma gondii* infection in wood mice (*Apodemus sylvaticus*) in Malham Tarn Nature Reserve …………………………..64

4.1. Introduction ………………………………………………………………………...65

4.2. Materials and Methods...………………………………………………………......…68

4.3 Results ……………………………………………………………………………….70

4.4 Discussion ……………………………………………………………………………...83

Chapter 5: Prevalence of infection and genotypic diversity of *Toxoplasma gondii* in patients with lung cancer ……………………………………………………………...89

5.1 Introduction ………………………………………………………………………....90

5.2 Materials and Methods ……………………………………………………………....94

5.3 Results ……………………………………………………………………………….95

5.3.1 Molecular analysis of the human lung tissue samples ……………………………….97

5.3.2 Immunohistochemistry (IHC) detection of *T. gondii* ……………………………....99

5.3.3 Haematoxylin and eosin staining …………………………………………………106

5.3.4 Summary of prevalence …………………………………………………………..107

5.3.5 Molecular analysis and RFLP genotyping ………………………………….…108

5.3.6 Scanning Electron Microscopy and Transmission Electron Microscopy ……….....127

5.4 Discussion ....…....128

Chapter 6: General Discussion ……………………………………………….....132

References 153

Appendix ……………………………………………………………………….....181

List of publications ……………………………………………………………………....194

List of Figures

Figure 1.1 A maximum parsimony tree for the isosporoid coccidia 2

Figure 1.2 Example of fluorescent protein labelling of subcellular organelles in *T. gondii* ….4

Figure 1.3 *Toxoplasma gondii* – the tachyzoite stage ……………………………………….…..5

Figure 1.4 The tissue cyst with bradyzoites………………………………………………….…..6

Figure 1.5 The life cycle of *Toxoplasma gondii* …………………………………………….…...7

Figure 1.6 The wood mouse (*Apodemus sylvaticus*) ……………………………………..….17

Figure 1.7 Ha Mire plantation ……………………………………………………………….…18

Figure 1.8 Tarn Woods ……………………………………………………………………......18

Figure 1.9 Tarn Fen …………………………………………………………………………….19

Figure 1.10 Traps set in Spiggot Hill ………………………………………………………..20

Figure 1.11 Spiggot Hill in winter ……………………………………………………………..20

Figure 2.1 Location of Malham Tarn within North Yorkshire…....…………………………...27

Figure 3.1 PCR amplification of *A. sylvaticus* loci ……………………………………………48

Figure 3.2 Four evolutionary clusters (K=4) inferred from STRUCTURE …………………50

Figure 3.3 Evaluation of the robustness of the grouping by STRUCTURE ………………...51

Figure 3.4 Principal Component Analysis (PCA) of four *A. sylvaticus* populations ………..52

Figure 3.5 The population structure of *Apodemus sylvaticus* …………….…………………56

Figure 3.6 Principal Component Analysis (PCA) of four *A. sylvaticus* populations (DIYABC) ……………………………………………………………………………………………..57

Figure 3.7 Historical scenario of *A. sylvaticus* populations …………………………….....58

Figure 4.1 Example of mammalian α-tubulin PCR of mice individuals ……………………70

Figure 4.2 An example of B1 PCR of Malham mice ………………………………………..71

Figure 4.3 SAG1 PCR of wood mice ………………………………………………………..71

Figure 4.3 Nested-PCR amplification of SAG2 5′ end locus ……………………………..…72

Figure 4.4 Agarose gel electrophoresis analysis of SAG2 PCR ……………………………73

Figure 4.5 PCR amplification of SAG3 locus ………………………………………………74

Figure 4.6 RFLP of SAG3 PCR products …………………………………….………..74

Figure 4.7 PCR amplification of GRA6 locus ……………………………………….....75

Figure 4.8 Seasonal changes in prevalence of *T. gondii* in wood mice ……………………..79

Figure 4.9 The prevalence of *T. gondii* in different age groups of wood mice ……….......79

Figure 4.10 The prevalence of *T. gondii* by locality and sex ……………………………......80

Figure 5.1 PCR amplification of the α-mammalian tubulin gen from lung tissue samples.....95

Figure 5.2 An example of B1 PCR gel electrophoresis on human lung cancer patients…......96

Figure 5.3 Example SAG1 PCR of lung tissue from human lung cancer patients…………...97

Figure 5.4 An example of nested-PCR amplification of SAG2 3 prime region 98

Figure 5.5 An example of nested-PCR amplification of *T. gondii* SAG2 5 prime region ……………………………………………………………………………………………..98

Figure 5.6 An example of PCR amplification of SAG3 locus from lung cancer patients…...99

Figure 5.7 Anti-*Toxoplasma gondii* antigen immunostaining of human lung and control tissues……………………………………………………………………………….....101

Figure 5.8 Haematoxylin and Eosin staining of human lung sections ……….………..…..106

Figure 5.9 A) PCR-RFLP analysis of the SAG2 5 prime region………………………….. 108

Figure 5.9 B) PCR-RFLP analysis of the SAG2 3 prime region…………………………...108

Figure 5.10 RFLP of SAG3 PCR products with *AlwnI* …………………………..….109

Figure 5.11 RFLP of SAG3 PCR products with *NciI* …………………………………...109

Figure 5.12 alt. SAG2 PCR and RFLP from the human lung tissue ………………………110

Figure 5.13 APICO PCR and RFLP from the human lung tissue …………………………111

Figure 5.14 BTUB PCR and RFLP from the human lung tissue ………………………….112

Figure 5.15 c22-8 PCR and RFLP from the human lung tissue……………………………113

Figure 5.16 c29-8 PCR and RFLP from the human lung tissue ...……………………..…...114

Figure 5.17 GRA6 PCR and RFLP from the human lung tissue...………………………...115

Figure 5.18 L358 PCR and RFLP from the human lung tissue ...…………………………116

Figure 5.19 PK1 PCR and RFLP from the human lung tissue………………………….…….117

Figure 5.20 Sequence analysis of the SAG3 marker (sample 833) ……………………..…121

Figure 5.21 Sequence analysis of the SAG3 marker (sample 1049) ………………………122

Figure 5.22 SEM and TEM of the lung tissue samples ………………………………….…….127

List of Tables

Table 2.2 Primer sequence of nPCR targeting the 5/ and 3/ end of SAG2 gene ……………31

Table 2.2 *A. sylvaticus* microsatellite primer sequences ……………………………………34

Table 2.3 Core repeats and PCR conditions for 10 microsatellite loci in *A. sylvaticus*……..35

Table 2.4 PCR conditions for *T. gondii* markers ……………………………………………42

Table 3.1 Microsatellite loci diversity for *Apodemus sylvaticus* ……………………………49

Table 3.2 Inference of genetic clusters (K) from *A. sylvaticus* …………………………...50

Table 3.3 Genetic differentiation among *A. sylvaticus* populations………………………….52

Table 3.4 Pairwise Population Matrix of Nei genetic distance …………………………...53

Table 3.5 Bottleneck results for four tested wood mouse populations …………………...…54

Table 3.6 Effective population size estimated with NeEstimator software …………………56

Table 3.7 The population structure of wood mice from the study site ……………………...57

Table 3.8 Table of Shannon Pairwise Analysis ……………………………………………..59

Table 4.1 PCR analysis of the brain tissue from *Apodemus sylvaticus*………...……………76

Table 4.2 Genotypes of *Toxoplasma gondii* in wood mice from the study site……………..77

Table 4.3 Prevalence of *T. gondii* infected *A. sylvaticus* at sites of collection………………80

Table 4.4 Prevalence of *T. gondii* infected *A. sylvaticus* with respect to genetic population..81

Table 4.5 Prevalence of *T. gondii* and genotypes present in 22 individual *A. sylvaticus*……82

Table 5.1 Frequencies of different stages of *T. gondii* detected with IHC in the lung tissue of cancer patients …………………………………………………………………………...103

Table 5.2 The matrix of successful PCR reproduction with *T. gondii* markers ……………119

Table 5.3 Distribution of *T. gondii* genotypes determined in lung cancer patients ………..120

Table 5.4 *T. gondii* genotypes determined in 88 human patients with lung cancer ………..123

Table 6.1 Seroprevalence of *T. gondii* detected among different cancer patients groups…..138

Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Professor Geoff Hide, for his support and guidance. I could not have asked for better role model, inspirational, supportive, and patient advisor. I would like to thank you for encouraging my research and for bringing me such an interesting problem to study.

I would like to give special thanks to Dr Robert Jehle, Dr Belgees Boufana, Dr Tony Bodell and Dr Jacqueline Hughes for their technical help, Michelle Hussein and Kellyanne Boyce for the sampling assistance, Helen Carlin and Dr. Lucy Smyth for the help with the human samples, Peter Shum for the help with the population softwares, Dr. Kevin Bown for the help with statistical analysis and Cath Hide for the help with histological techniques.

I would like to thank you The National Trust for granting the license which enabled this work.

A special thanks to my partner Alessandro, family and friends.

ABBREVIATIONS

Complement Fixation Test (CF)

Chronic obstruction pulmonary disease (COPD)

Diaminobenzidine (DAB)

Dye Test (DT)

Enzyme Linked Immunoabsorbent Assay (ELISA)

Fishers exact test (FET)

Glycosylphosphatidylinositol (GPI)

Healthy ex-smoker (HEX)

Healthy non-smoker (HNS)

Healthy smoker (HS)

Haematoxylin eosin staining (HE)

Immunohistochemistry (IHC)

Immunoglobulin G (IgG)

Immunoglobulin M (IgM)

Indirect Haemaglutination Test (IHA)

Indirect Fluorescent Antibody Tests (IFA)

Infinite allele model (IAM)

Interferon γ (IFN-γ)

Markov Chain Monte Carlo (MCMC)

Micro RNA (miRNA)

Modified Agglutination Test (MAT)

Non-small lung cancer (NSCLC)

Polymerase chain reaction (PCR)

Principal Component Analysis (PCA)

Restriction Fragment Length Polymorphism (RFLP)

Ribonucleic acid (RNA)

SAG1 related sequence (SRS)

Scanning electron microscopy (SEM)

Short tandem repeat (STR)

Small cell lung cancer (SCLC)

Stepwise mutation model (SMM)

Surface antigen gen (SAG)

Transmission electron microscopy (TEM)

Variable number tandem repeat (VNTR)

Declaration

I hereby declare that the work that is presented in this thesis is my own work unless otherwise stated. Details are given in each chapter where any of the results are produced in collaboration with other members of the research group. I declare that I have not submitted this work for a degree or any other qualification at this or any other university.

…………………………

Jaroslav Bajnok

January 2017

Abstract

Toxoplasma gondii is a globally distributed protozoan parasite that infects humans and a wide variety of warm-blooded animals. Although there are many surveys for *T. gondii* infection in mammals, little is known about the detailed distribution in localised natural populations. In this study we investigated host genotype and spatial location in relation to *T. gondii* infection and genotypes. We collected wood mice (*Apodemus sylvaticus*) from 4 sampling sites within a localised peri-aquatic woodland ecosystem which is relatively free of cats. Mice were genotyped using standard *A. sylvaticus* microsatellite markers and *T. gondii* and its genotypes were detected using 5 specific PCR based markers: SAG1, SAG2, SAG3, B1 and GRA6 directly from infected tissue. Of 126 wood mice collected, 44 samples gave positive reactions with *T. gondii* specific markers giving an infection rate of 34.92% (95% CI: 27.14%-43.59%). A total of 24/76 (31.58%, 95% CI: 22.19%-42.74%) males and 20/50 (40%, 95% CI: 27.59%-53.84%) female mice were found to be positive for *T. gondii* with no significant difference (P = 0.353). Juvenile, young adults and adults were infected at similar prevalences, respectively, 7/17 (41.18%), 27/65 (41.54%) and 10/44 (22.72%) with no significant age-prevalence effect (P = 0.23). Detailed analysis of the RFLP patterns and the DNA sequences for the SAG2, SAG3 and GRA6 loci showed a range of genotypes but, surprisingly, suggested that 30/44 (68.2%) infected mice had multiple genotypes (mixed infections) present. Results of genetic analysis of the mice showed that the collection consists of four genetically distinct populations. There was a significant difference in *T. gondii* infection in the different mouse genotypically derived populations (P=0.035) but not between geographically defined populations based on sampling location (P=0.29). In a parallel study, DNA was successfully collected from 88 human lung tissue samples. All samples showed successful amplification of the human α-tubulin gene and were used for *T. gondii* DNA detection. We used commonly used PCR markers (B1, SAG1, SAG2, SAG3, GRA6, APICO, L358, PK1, SAG1-Su, BTUB, alt.SAG2, c22-8 and c29-2), histological and immunohistochemical staining to confirm the presence of the parasite. All 88 tested samples were confirmed to be positive for *T. gondii* with markers B1, SAG1, SAG2 3', SAG2 5' and SAG3, giving a prevalence of 100% (95% CI: 95.82%-100%). From all successfully genotyped samples, 34 had single infection on all loci and 42 were of mixed infection on one or more loci with all three genotypes present. Type II genotype was the most predominant, followed by Type I and Type III. We detected 11 unusual genotypes. Immunohistochemistry

was performed on 76 of the 88 tissue sections using commercial polyclonal antibodies produced in rabbits. All 76 sections were confirmed to be positive for *T. gondii.* A surprisingly high number of patients (96.05%) showed evidence of an active form of infection, as defined by the presence of tachyzoites or infected alveolar macrophages (or other cell types). Only three subjects (3.95%) had the dormant cyst stage as the only stage present. All 76 tissue sections were successfully stained with haematoxylin and eosin and observed under the light microscope. The presence of structures consistent with infection by the parasite was confirmed in 67 samples. All these patients are at risk of reactivation of chronic infection, leading to toxoplasmic encephalitis or pulmonary toxoplasmosis; which can complicate and delay their treatment or lead to death.

Chapter 1:

General Introduction

1. Introduction

Toxoplasma gondii is an obligate intracellular parasite capable of infecting any warm-blooded animal including humans (Black *et al*. 2000, Dubey *et al*. 2001). Its distribution is global, ranging from Alaska to Australia and it infects up to a third of human world's population (Jackson and Hutchinson, 1989). It is more prevalent in the temperate and moist tropical regions of the world than in the arctic zones or in the hot arid regions. It is unique in that it infects any type of nucleated cells (Dubey *et al*. 1988). It has been found in intestinal epithelium, hepatocytes, alveolar cells, macrophages, retinal cells, brain cells and in red blood cells of birds. This protozoan parasite was discovered in 1908 by Nicolle and Manceaux in Tunisia in rodents *Ctenodactylus gundi.* At about the same time Splendor (1909) found it in rabbits in Brazil. In 1923, Janků described the *T. gondii cyst* in the retina of hydrocephalic child, but the role of the parasite as a human pathogen was not known until Wolf and Cowen (1942) identified toxoplasmosis as a disease of humans. The name *Toxoplasma* (toxon = arc, plasma = form, Greek) is derived from its crescent shape.

Toxoplasma gondii is a protozoan parasite belonging to the Phylum Apicomplexa (which contains many human and veterinary disease agents (e.g. *Plasmodium* spp., *Babesia* spp., *Eimeria* spp., *Sarcoscystis* spp.). Class Conoidasida, Subclass Coccidiasina, Order Eucoccidiorida, Family Sarcocystidae, Subfamily Toxoplasmatinae.

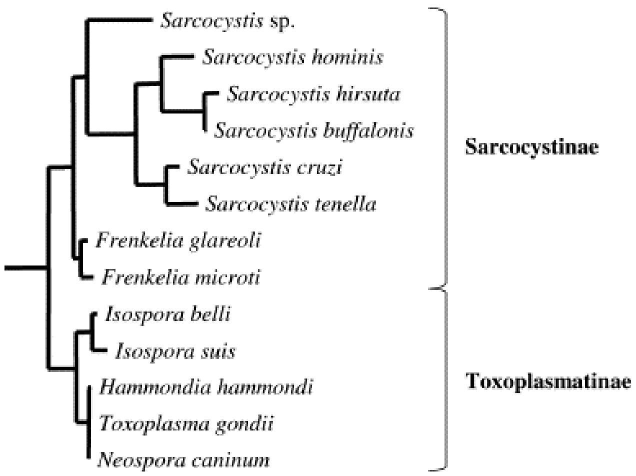


Figure 1.1 A maximum parsimony tree for the isosporoid coccidia (a single branch of the tree for the entire phylum Apicomplexa) that corresponds to the family Sarcocystidae. The current sub-families Sarcocystinae and Toxoplasmatinae are reasonably well-supported in this reconstruction based on the 18S ribosomal RNA sequences (Barta *et al*. 2001).

The full life cycle was not known until 1970 when several groups independently discovered the oocyst stage in cat faeces (Hutchison *et al.* 1969, 1970, 1971; Frenkel *et* al. 1970; Dubey *et al.* 1970; Sheffield and Melton, 1970; Overdulve, 1970; Weiland and Kuhn, 1970; Witte and Piekarski, 1970). The discovery of *T. gondii* oocysts in cat faeces and its implications have been reviewed by Frenkel (1970, 1973).

Despite high prevalences in some areas, few studies have been conducted on natural populations of animals. Most of the tested samples were collected from domestic animals in farms, slaughterhouses, processed meat and human studies. The majority of epidemiological studies in wild animals have been based on serological diagnostic methods or isolation of the parasite. Only in few studies have the parasite bee investigated in an individual host population. For example, Dodd *et al.* (2014) used microsatellites to investigate the population of bats infected with *T. gondii* but to our knowledge no similar study has ever been conducted on wild rodents.

1.1. Life cycle and morphology

The definitive hosts of *T. gondii* are domestic cats and a number of other members of the family Felidae. It is only in these animals that the sexual stages of the parasite occur. There are three infectious stages of *T. gondii:* tachyzoites (in pseudocysts or clones), bradyzoites (in tissue cysts) and sporozoites (in oocysts). These individual parasite stages possess the apical complex which is used for cell invasion. The apical complex consists of a set of spirally arranged microtubules (the conoid), a secretory body (the rhoptry) and one or more polar rings (Dubey *et al.* 1988). Additional slender electron dense secretory bodies (micronemes) surrounded by one or two polar rings may also be present.

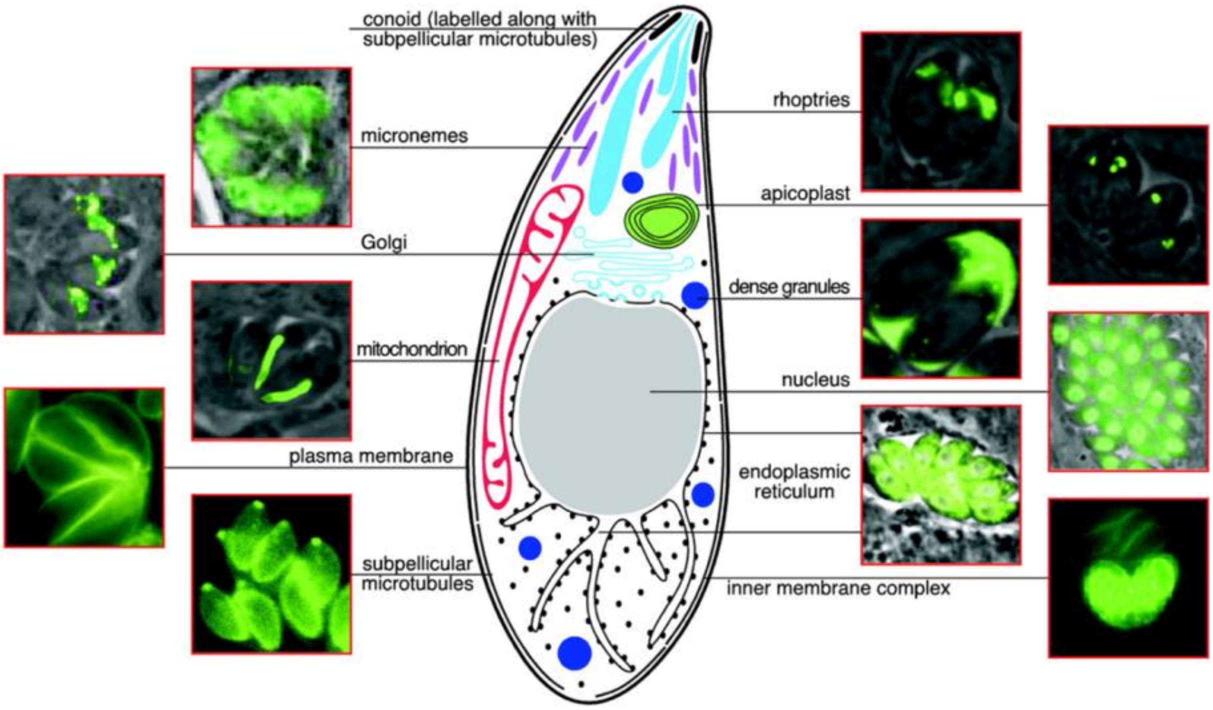


Figure 1.2 Example of fluorescent protein labelling of subcellular organelles in *T. gondii*. (Joiner and Roos, 2002).

Tachyzoites **(**Greek. *tachys*, fast + *zoon*, animal)

The tachyzoite is often crescent shaped and is approximately 2 x 6 µm. The nucleus is usually situated toward the posterior end or in the central area of the cell. This is the rapidly dividing stage that can multiply in any nucleated cells of the definitive and intermediate host. They have no cilia or pseudopodia but they are able to rotate, glide and undulate (Chiappino, *et al.* 1984). Tachyzoites actively penetrate the host cell wall or are internalised during phagocytosis (e.g. macrophages) and becomes surrounded by a parasitophorous vacuole, inside which endogeny occurs. During penetration parasites orientate themselves so that the apical end is facing the host cell. The micronemes are expelled first and expulsion occurs with initial contact between parasite and host cell. The rhoptries are discharged immediately after the micronemes, and dense granule contents are released last. Many dense granules are extruded after the parasite has completed its entry, and thus probably play a role in modifying the host cell (Baum *et al.* 2008). The tachyzoites multiply asexually by repeated division (endogeny, schizogony). Host cells containing numerous tachyzoites are called clones (terminal colony, pseudocysts). After a few divisions they finally rupture; releasing 5-32 tachyzoites (Dubey; Frenkel, 1972) with each tachyzoite of invading another host cell. Spread can be local, across serous cavities, by the blood stream or via the lymphatic system. At least five generations of such propagative stages have been described (Dubey, 1972). Invasions and spread continue until the host dies or immunity is established.

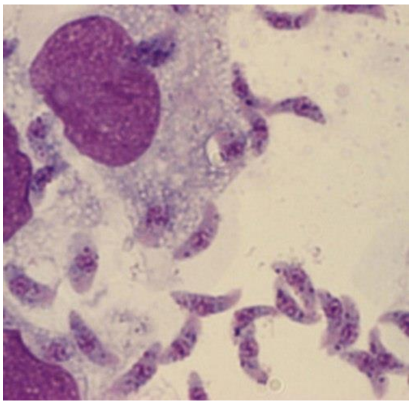


Figure 1.3 *Toxoplasma gondii* – the tachyzoite stage, Giemsa stain, (x100) (http://www.studyblue.com/)

Bradyzoites (Greek. *bradys*, slow + *zoon*, animal)

After a few divisions *T. gondii* forms tissue cysts that vary in size. The cyst wall is elastic, thin (<0.5μm) and may enclose hundreds of crescent – shaped, slender *T. gondii* organisms known as bradyzoites. Young tissue cysts may be as small as 5 μm in diameter and contain only two bradyzoites, while older ones may contain hundreds of organisms. Tissue cysts in the brain are often spheroidal and rarely reach a diameter of 70 μm, whereas intramuscular cysts are elongated and may be 100 μm long (Dubey, 1977). Bradyzoites are 7 x 1.5 μm in size and differ only slightly in morphology from tachyzoites. They are smaller and the nucleus is situated towards the posterior end (Dubey, 1998). The cyst wall is ultimately lined by granular material which also fills the space between bradyzoites. It is likely that the host cells contribute to this deposition, because in electron photomicrographs of developing cysts a heavy concentration of mitochondria and host cell endoplasmatic reticulum lines the outside of the parasitophorous vacuole. Bradyzoites are less susceptible to destruction by proteolytic enzymes than tachyzoites. Tissue cysts may occur in many organs, including liver and kidneys; they are more common in neural and muscular tissue, particularly the brain, eye and may persist for the life time of the host. Tissue cysts are more common in the latent stage of infection after the host has acquired immunity. The cyst form is well adapted to transmission to carnivores, because the bradyzoites released from the cyst when it is exposed to gastric juices are able to survive for at least 3 hours (Jacobs *et al.* 1960).

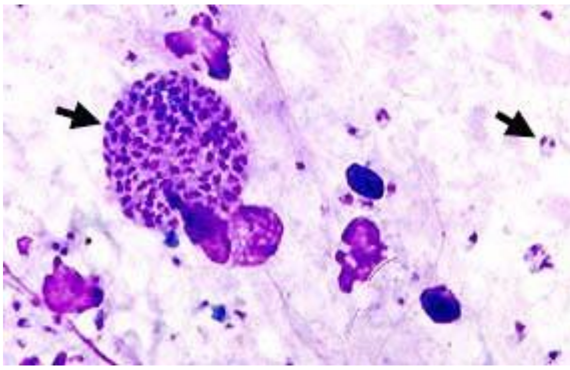


Figure 1.4 The tissue cyst with bradyzoites, Giemsa stain, (x100) (http://pathhsw5m54.ucsf.edu/overview/toxo.html)

Oocyst

Oocysts are the environmentally resistant stage in the life cycle of coccidia. Oocysts of *T. gondii* are formed only in felids. After the ingestion of tissue cysts by cats, the tissue cyst wall is dissolved by proteolytic enzymes in the stomach and small intestine. The released bradyzoites penetrate the epithelial cells of the small intestine and initiate development of numerous generations of asexual and sexual cycles (Dubey and Frenkel, 1972). Microgametogony occurs in manner similar to schizogony. Fertilization of the macrogamete by the microgamete results in the formation of a zygote which measures 10 x 12 µm. The zygote is unsporulated when shed, but the sporulation takes place in outside environment and takes around 1- 5 days depending on aeration and temperature (Frenkel and Dubey, 1972). The sporoblast gives rise to two sporocysts each containing four sporozoites. The sporozoites are 2 x 6(8) µm in size. The sporulated oocysts are highly infective and can remain viable in the environment for many months (Lass *et al.* 2009).

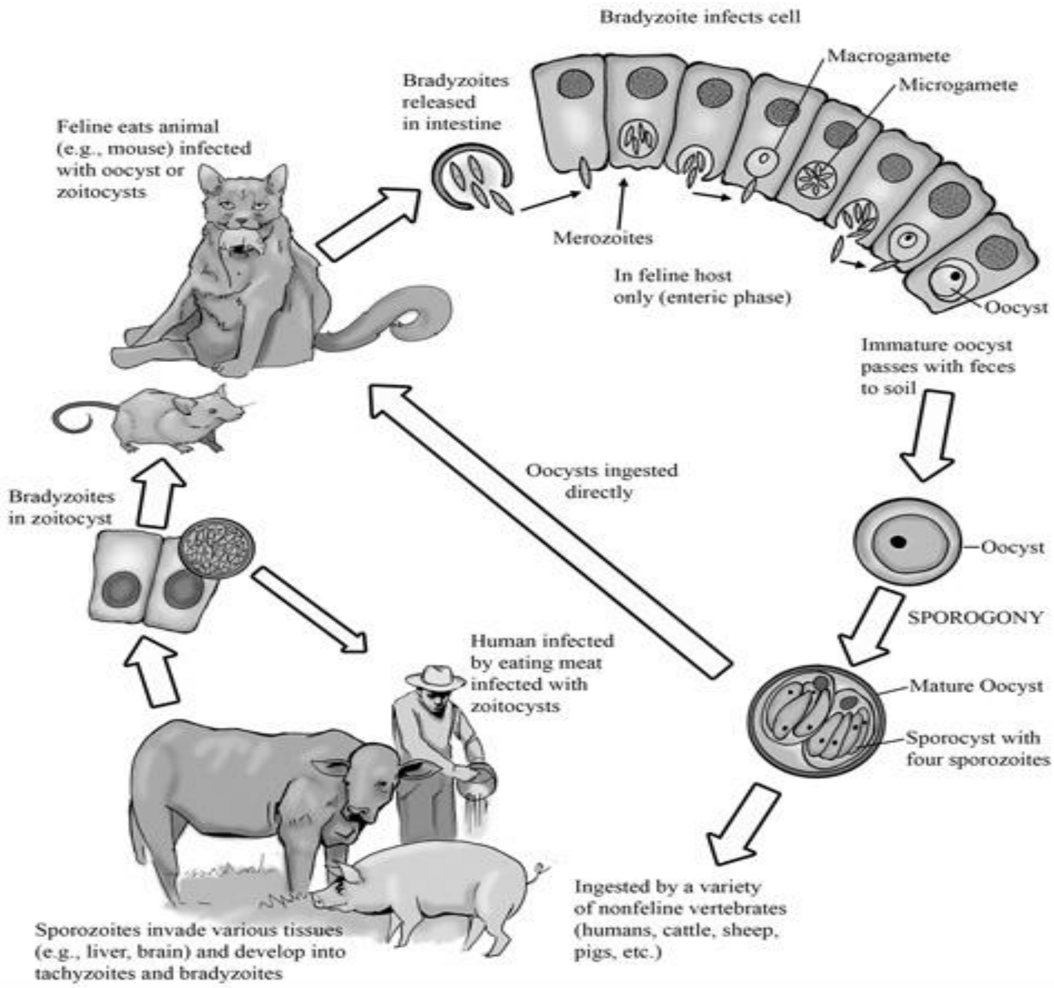


Figure 1.5 The life cycle of *Toxoplasma gondii*

(http://clem.mscd.edu/~churchcy/BIO3270/Images/Protozoans/Toxoplasma_gondii.htm)

The only known definitive hosts for *T. gondii* are members of family Felidae (domestic cats and their relatives), where the sexual stage occur within the intestinal epithelium. The domestic cat seems to be the true host of *T. gondii.* The jaguarondi and ocelot, bobcat, Asian leopard cat and mountain lion have been experimentally infected with tissue cysts and have shed oocysts (Jewel, Miller, 1972). The definitive host becomes infected by ingesting infected animals or sporulated oocysts. Prepatent periods to the shedding of oocysts vary with the stage of *T. gondii* ingested: 3 to 10 days after ingesting tissue cysts, 19 days or longer after ingesting tachyzoites, and 20 days or longer after ingesting oocysts (Wallace, 1973, Dubey, 1977). Unsporulated oocysts are shed in the cat's faeces. Oocysts are usually shed only for 1-2 weeks but large numbers may be shed. Oocysts take 1-5 days to sporulate in the environment and become infective. Intermediate hosts in nature (mammals, birds) become infected after ingesting soil, water or plant material contaminated with oocysts. Poikilothermic vertebrates can also be infected if their body temperature is raised high enough. Ingested oocysts release the sporozoites, and they transform into rapidly dividing tachyzoites. At this stage of the infection congenital transmission may occur as tachyzoites can cross the placenta. The tachyzoites localized in neural and muscle tissue usually develop into tissue cyst bradyzoites (Dubey *et al.* 1976). Cats become infected after consuming intermediate hosts harbouring tissue cysts. Cats may also become infected directly by ingestion of sporulated oocysts. The bradyzoites tissue cysts may be ingested by another intermediate host through carnivory.

1.2.Transmission

Most apicomplexan genera consist of numerous species each of which interacts with a specific host. *Toxoplasma gondii* has successfully colonized a diverse range of host species from very different environments. The most general mode of spread of toxoplasmosis depends on the presence of the oocysts, which are produced only by the intestinal stages in the cat and the other members of the family Felidae. The proportion of cats excreting oocysts at any one moment is not high, being usually not more than 2% (Dubey, 1977). Oocysts are shed for only a short period of 1-2 weeks in the life of the cat; before the body's immune response stops oocyst production altogether, however the high numbers shed assure widespread contamination of the environment. It is rare for a cat to re-shed oocysts in faeces after their first infection, and when this does occur it usually results in a much smaller number of oocysts. Cats can also be infected by ingesting oocysts, but experiments have shown that these are much less potent sources of infection for cats than tissue cysts (Dubey, *et al.* 1976). Oocysts are highly infective to mice; however mice may not become infected following ingestion of up to 100 bradyzoites (Dubey, 2005). Depending on humidity, the oocysts in the soil persist for weeks or months to a year and a half (Frenkel, 1973). The feline habit of covering their faeces enhances the survival of oocysts since the humidity is greater underground than on the surface. Cold inhibits sporulation of oocysts and unsporulated oocysts are killed in from 1 to 7 days of constant freezing. Sporulated oocyst can survive constant freezing at -20˚C for 28 days (Hutchinson, 1967). Oocysts can be mechanically transferred by many animals on their surfaces, flies, cockroaches, earthworms and by weather conditions such as rain and snow (Chinchilla, 1976, Wallace 1972). Transmission by oocysts is the only mode of horizontal transmission to herbivores and one of the modes of transmission to omnivores and carnivores. Toxoplasmosis can be perpetuated among carnivorous and omnivorous animals by cannibalism or by poor animal husbandry practices. In recent years, in many countries, a number of cases where toxoplasmosis has been contracted via a water source have been reported. In Panama in 1979 *T. gondii* infection was observed among soldiers stationed in the jungle. Epidemiological investigation showed that the source of infection was water from the stream contaminated with oocysts excreted by jungle cats (Benenson, 1982). Cases of toxoplasmosis were also confirmed among a group of vegetarians in India, which were related to the contamination of drinking water with the oocysts (Hall *et al.* 1999).

The infection in humans and carnivores is probably most often the result of ingestion of tissue cysts contained in undercooked meat (Dubey and Beattie, 1988; Roghmann *et al.* 1999; Lopez *et al.* 2000), though the exact contribution of food-borne toxoplasmosis versus oocyst-induced toxoplasmosis to human infection is currently unknown. Viable tissue cysts have been found in tissues of naturally infected animals, particularly in pigs and sheep, and also in wild game. They are rare in cattle. Some studies suggested that tissue cysts are killed by commercial procedures of curing with salt, sucrose, or low temperature smoking (Lundén *et al.* 1992). Therefore, it has previously been suggested that processed meat is an unlikely source of infection for humans.

Prevalence of infection varies considerably, both according to country and author. In the course of evolution, *T. gondii* has developed a broad range of potential routes of transmission. However, the elucidation of these routes during the past three decades has not elucidated which of these routes is more important epidemiologically.

Over the last 30 years, the incidence of prenatal infection with *T. gondii* has been estimated to vary from 1 to 100 per 10 000 births in different countries. The risk of intrauterine infection of the foetus, the risk of manifestation of congenital toxoplasmosis, and the severity of the disease depend on the time of maternal infection during pregnancy, the immunological competence of the mother during parasitaemia, the number and virulence of the parasites transmitted to the foetus, and the age of the foetus at the time of transmission (Tenter *et al.* 2000).

1.3 Human toxoplasmosis

Toxoplasma gondii infects all warm-blooded animals including humans (Black *et al*. 2000). It is estimated that 30 % of the human population is infected (Wong and Remington, 1993; Petersen and Dubey, 2001). The prevalence can be between 15 % up to 85% depending on geographical location (Dubey and Beattie, 1988). *Toxoplasma gondii* is the third most common food contracted disease in the United States, where approximately 225 000 cases are reported each year (Jones, 2001). Toxoplasmosis is more common in developing countries and in lower socioeconomic communities (Barry *et* al. 2013) but, for example around 50% of the adult population in Germany is infected and as many as 90% of adults in Paris are seropositive (Hökelek *et al.* 2013). High seroprevalence (59.8%) is documented in Inuit populations of Nunavik and other northern communities (Messier *et al.* 2009). Humans usually get infected when they consume undercooked meat, through contaminated food (vegetables), water, in hospitals if they get infected transplants, congenitally or directly with oocysts from the cat faeces. Infection in immunocompetent individuals is usually asymptomatic but toxoplasmosis can establish itself as lifelong latent infection. In this form of disease cysts are formed and survive in various tissues of the body, including brain. *Toxoplasma gondii* infection can cause severe disease in immunocompromised patients (AIDS, transplant and oncological patients) and in the foetus of mothers who are infected during pregnancy (Dubey *et al.* 2008). In these patients the disease transits into very acute phase and if left untreated the patients usually die of encephalitis. Other symptoms are myocarditis, toxoplasmic pneumonitis, fever, cough, dyspnoea and retinochoroiditis. In France, 37% of patients with AIDS have evidence of toxoplasmic encephalitis (Hökelek *et al.* 2013). The reports show that most of the cases of toxoplasmosis in immunocompromised patients are a consequence of latent infection and reactivation (Porter *et al.* 1992).

Ocular toxoplasmosis is usually a self-limiting condition in immunocompetent individuals. In previous reports, 7%-33% of all European patients and 27%-92% of American patients developed an ocular lesion (Freeman *et al.* 2008). Retinochoroiditis usually results from reactivation of congenital infection, although cases have been reported that were part of acute infection (Freeman *et al.* 2008). Classic treatment of ocular toxoplasmosis has been with pyrimethamine, sulfadiazine and prednisone for 4 to 12 weeks (Rothova *et al.* 1993).

Recent reports suggest that latent toxoplasmosis can play an important role in many psychiatric and neurological disorders such as bipolar disorder, personality disorder, Parkinson disease (Miman *et al*. 2010), Alzheimer disease (Kusbeci *et al.* 2011), epilepsy

(Yazar *et al.* 2003), schizophrenia (Torrey and Yolken, 2005; Niebuhr *et al.* 2008), autism and brain tumours (Vittecoq *et al.* 2012). The incidence of suicides in Europe correlates with prevalence of toxoplasmosis in particular country (Ling *et al.* 2011). *Toxoplasma* also increases the reaction time in infected subjects, which could possibly explain the increased probability of being involved in a traffic road accident (Flegr, 2002). All these behavioural changes are thought to be caused by bradyzoite cysts in the brain and their production of tyrosine hydroxylase, the enzyme necessary for the production of dopamine (Prandovzsky *et al.* 2011).

The most significant manifestation of toxoplasmosis in the foetus is encephalomyelitis. *Toxoplasma gondii* infection often leads to abortion or severe deformation of the foetus (hydrocephalus). Many infants develop clinical symptoms later during their lives. Multiple complications can occur in persons with congenital toxoplasmosis, including mental retardation, seizures, blindness and deafness. Women with latent toxoplasmosis have a higher probability to have a male offspring and there is an increased risk of giving birth to a child with Down syndrome (Hostomska *et al.* 1957). Approximately 10-20 % of pregnant infected women become symptomatic (Montoya *et al.* 1996). Routine serological screening should be performed on all pregnant women considered to be at risk for primary *T. gondii* infection. Amniocentesis should be offered to identify infection in amniotic fluid by PCR.

Immunoglobulin M (IgM) is the first antibody subclass to appear in the human immune response. Immunoglobulin G (IgG) is increased in the secondary response. One study found that IgM remained raised in patients for 12 to 18 months, while IgG activity stayed low for a maximum of 4 months (Kodym *et a*l. 2007). If IgG and IgM are both negative, this can indicate the absence of infection or recent acute infection. If both are positive, this indicates either a recent infection or a false –positive test result. Interferon-γ (IFN-γ) secreted in response to *T. gondii* infection plays essential role in maintaining toxoplasmosis in latent form. IFN-γ is reduced in schizophrenia patients but elevated in the latent toxoplasmosis (Suzuki *et al.* 2005).

1.3.1 Diagnostics

Toxoplasmosis is generally benign and has little clinical relevance in immunocompetent patients. However it can be serious or fatal in children infected during their intra-uterine development and in immunocompromised patients. Parasitaemia occurs early and it is probable that all organs become infected. Histological changes can be found in many but not all organs in the majority of host species. They occur very rarely in kidneys of any species and very rarely in any organs of rats (Soulsby, 1978). Diagnosis is made by biological, serological, or histological methods, or by some combination of them. Clinical signs of toxoplasmosis are nonspecific and cannot be depended upon for a definitive diagnosis. *Toxoplasma* cysts do not provoke inflammatory reaction and only rarely are they found near one, even when they are, they still appear intact and unchanged by the inflammation unless there has been extensive necrosis. A rapid diagnosis may be made by microscopic examination of impression smears stained with Giemsa stain. Well – preserved tachyzoites are crescent shaped and stain well with any Romanowsky stain. Tissue sections stain well with haematoxylin and eosin (HE). Histological methods are relatively cheap and can confirm diagnosis within few hours but they have been proved insensitive by some studies (Garcia *et al.* 2006; Esteban-Redondo *et al.* 1999; da Silva and Langoni, 2001). The finding of tachyzoites indicates active infection and is diagnostic; finding of tissue cysts may indicate latent infection and is not necessarily diagnostic of disease.

Titres of *T. gondii* antibody rise rapidly to a maximum in 4-6 weeks, are maintained at this level for a few weeks and then fall steadily. Many serologic tests have been used for the detection of IgG and IgM *T. gondii* antibodies.

Dye Test (DT)

The dye test requires human serum without antibody as accessory factor, but has an added disadvantage of requiring the use of live organisms. The tachyzoites are incubated with accessory factor and the test serum at 37˚C for 1hour and a dye (methylene blue) is added. Tachyzoites unaffected by antibody are stained uniformly with methylene blue. Specific antibody induces complement mediated cytolysis of tachyzoites and the cytoplasm leaks out (Endo, 1976; Schreiber, 1980). As a result tachyzoites affected by antibody do not incorporate methylene blue. This test is highly specific and sensitive with no evidence of false results in human. Although DT measures antibodies specific to *T. gondii* in many hosts, ruminant sera have substances that give false results unless sera are inactivated at 60˚C for 30 min (Berger *et al.* 1966).

Indirect Haemaglutination Test (IHA)

Soluble antigen from tachyzoites is coated on red blood cells which are then agglutinated by immune serum. The IHA test is frequently negative in congenital infections and does not become positive as early as DT.

Complement Fixation Test (CF)

It is generally believed that complement antibodies appear later than DT antibodies and the test is positive during acute infection, but this largely depends on the antigenic variation.

Indirect Fluorescent Antibody Test (IFA)

The IFA test is sensitive and specific, but requires the appropriate fluorescein labeled antispecies globulin free of antibody to *T. gondii*. In the conventional IFA test, whole killed tachyzoites are incubated with serum and antibody detection is enhanced by adding fluorescent-labeled antispecies IgG (or whole immunoglobulin) and viewing with a fluorescent microscope. A modification of IFA (IgM-IFA) was developed (Naot and Remington, 1980). The main disadvantage is the need of a fluorescent microscope and species specific conjugates and the presence of non-specific cross reaction with rheumatoid factor and antinuclear antibodies. Because IgM is not placentally transferred, the IgM-IFA test demonstrates the antibody developed by an infant, indicating infection rather than passively transferred antibody. Latent toxoplasmosis is typically accompanied by a stable low antibody titer; retinochoroiditis (Jacobs, 1954). In reactivated toxoplasmosis, the antibody titer tends to rise probably because antibody forming B-cells are less affected by immunosuppression than T-cells mediating cellular immunity.

Enzyme-Linked Immunoabsorbent Assay (ELISA)

In the ELISA test, soluble antigen is absorbed on a plastic surface (microtiter plates or slides) and the antigen-antibody reaction is enhanced by the addition of a secondary enzyme-linked antibody–antigen system, and the reaction can be assessed objectively by quantitation of the colour that develops. Numerous modifications (Schaefer *et al.* 2011) of ELISA have been published (Immunoglobulin M Immunoabsorbent Agglutination Assay Test).

More methods can be used for the detection of the parasite, but they are not commonly used in routine tests (Antigen-Specific Lymphocyte Transformation, Skin test, Immunoperoxidase staining, Detection of Circulating Antigens).

Serological methods have often poor efficiency, especially in neonates and in immunodeficient patients. To improve these diagnoses there have been many studies with various immunological assays to detect the parasite or its components either in experimental infection or in human toxoplasmosis (Raizman, 1975; Hafid, 1995).

Polymerase Chain Reaction detection (PCR) of *T. gondii*

PCR assays for the detection of *T. gondii* DNA were developed and primers specific to different regions of the genome of the parasite have been used. These assays demonstrated high specificity and sensitivity in various human samples (blood, tissue and amniotic fluid). Hafid (2001) performed a study where PCR was compared with capture ELISA and immunoblotting for the detection of *T. gondii* in sera of acutely infected mice. PCR (B1 gene) was the most sensitive assay and detected the parasite from 18 h post infection, whereas the other assays detected it only after 24h. However the sensitivity of the results depends on the strain of *T. gondii*, the load of the inoculum and the PCR protocols used. Nested PCRs followed by hybridization are the most sensitive; however they are not quantitative and are time-consuming. Using the marker for the B1 gene we can detect as little as 50fg of *T. gondii* DNA which is the equivalent of single tachyzoite (Jones *et al.* 2000). Other commonly used markers target the P30 gene, which encodes highly immunogenic major surface proteins - SAGs (Rodriguez *et al*. 1985). Most of the *Toxoplasma* surface antigens belong to the developmentally regulated and distantly related SAG1 or SAG2 families. The genes encoding the surface antigens are distributed throughout the *T. gondii* genome, with remarkably little polymorphism being observed at each locus. The surfaces of tachyzoites and bradyzoites are covered with glycosylphosphatidylinositol (GPI)-anchored antigens (Nagel *et al*. 1989) most of which are members of the surface antigen 1 (SAG1) or SAG2 families (Lekutis *et al.* 2001). Collectively, these surface antigens are known as the SRS (SAG1-related sequences) superfamily of proteins. SAG1 antigen contains many cysteine residues which lead to disulphide bridge formation. Paralogous SRS sequences are 24-99% identical, are tandemly arrayed throughout the genome, and are present on most, if not all chromosomes. Genotypic differences among SRS sequences are present at several loci (e.g. the absence of SAG5B, the truncation of SAG2D in Me49 compared with RH (Jung *et al.* 2004). These molecules appear to play a role in host cell invasion, immune modulation and virulence attenuation, although they may also provide protection needed by the parasite to survive in the environment. It is possible that SAG1-related proteins provide a structural barrier required by the parasites to survive inside mammalian hosts. Alternatively, the major function of the SAG proteins may be as immune modulators or virulence factors, since the survival of the parasite is dependent upon the survival of the host. Thus, it could be argued that the high level of expression of SAG1, which is limited to the tachyzoite, could assist the host immune response in tachyzoite elimination (Lekutis *et al.* 2001).

More than 50 different genetic markers are available for *T. gondii* isolate typing. Multilocus genotyping is used to reveal atypical or recombinant genotypes. Quantification of *T. gondii* by qPCR has been applied to the detection of the parasite in animal tissues, whole blood and amniotic fluid. Real-time PCR can detect low concentrations of target DNA and quantify starting copies of specific template. This PCR has been successfully used to detect *T. gondii* in human blood, cerebrospinal fluid, amniotic fluid and other samples. *Toxoplasma gondii* has a highly clonal population structure comprised of three widespread genotypes referred to as type I, type II and type III (Dardé, 1988; Howe, 1997). In humans, type II largely predominates in congenital toxoplasmosis, at least in Europe, but type I, atypical and recombinant genotypes are more frequently associated with severe forms of toxoplasmosis (Ajzenberg, 2002). In Europe, genotypes different from clonal type II were found in Spain, Poland, Portugal and Germany in chickens, pigs and pigeons (de Sousa *et al.* 2006; Waap *et al.* 2008; Berger-Schoch *et al.* 2011). All of the 3 lineages can be found in isolates from humans, however the majority fall into the type II genotype. Recently a fourth clonal type (haplogroup 12) has been discovered in North America (Khan *et al.* 2011). The clonal lineages found in the *T. gondii* population structure are closely related to virulence in human and mice (Howe and Sibley, 1995). The three clonal lineages *T. gondii* system is based on the results of SAG2 typing and only recently after new multilocus typing was used more unusual strains have started to emerge. Today, 11 *T. gondii* markers (SAG1, SAG2, SAG3, GRA6, APICO, L358, PK1, and BTUB, alt.SAG2, c22-8 and c29-2) are commonly used to identify the isolate (Su *et al.* 2010). Up to now 231 strains are known and their RFLP genotypes are available in the database (www.toxodb.org). Microsatellite typing is now considered one of the most powerful tools to study DNA polymorphism and is often used in combination with RFLP markers. A total of 15 microsatellite markers (TUB2, W35, TgM-A, B17, B18, M33, IV.1, XI.1, M48, M102, N60, N82, AA, N61 and N83) have been used to genotype *T. gondii*. Emerging new clonal types may show different levels of virulence in intermediate hosts.

1.4 The ecology of the wood mouse (*Apodemus sylvaticus*)

Rodents are a very suitable intermediate hosts for *T. gondii* transmission as felids often prey on them. Experimental infections of mice with infective oocysts showed a high frequency of vertical transmission from one generation to the next (Owen and Trees, 1998; Marshall *et al.* 2004) which suggests that mice could be a potential reservoir of *T. gondii* in natural populations. The wood mouse (*A. sylvaticus*) is a widespread wild rodent and is commonly found throughout Europe (except northern Scandinavia and Finland), British Isles and nearby islands (Nowak, 1991). It is an inhabitant of mainly woodland, grassy fields and cultivated areas, but is highly adaptable and is found in most habitats. Most wood mice build nests in underground burrows, tunnels, inside hollow logs or in dense vegetation. Males normally inhabit an area of approximately 109 m in diameter and females live in an area of 64 m in diameter. Most wood mice stay in the same area but can travel 400 m in one night (Nowak, 1991). They are mainly active during the dark, and are very good climbers. The breeding season begins in March and continues through until early winter if it is mild and there is a good food supply. Breeding males range over larger areas occupied by a number of females which produce up to four litters annually, with four to seven young each litter. The maximum life span is approximately 18 - 20 months, but juvenile survival is negatively affected by the presence of adult males (Bengston, 1989). Adult males are aggressive to one another and are driven from the nest soon after weaning. Many wood mice do not live more than four months and the overall population has an annual variation, with the highest peak in autumn and lowest in the spring. Wood mice are primarily granivores and their diet consists of grains, seeds, nuts, roots, berries, fruits, small snails and insects (Nowak, 1991; Parker, 1990). During the winter they do not hibernate but group together when sleeping. During severe winter seasons they fall into a sort of torpor, with a decrease in physiological activity. Wood mice are seen as pests and are responsible for damage to seedlings in forests and crops (Nowak, 1991).

Figure 1.6 The wood mouse (*Apodemus sylvaticus*)

1.5 Malham Tarn

Malham Tarn is situated in the Yorkshire Dales, a National Park in the Yorkshire Pennines, approximately 25 miles (40 km) north-west of Bradford and about 2.5 miles (4.0 km) north of the nearest settlement, Malham. The glacial lake is 377 metres (1,237 ft.) above sea level, which makes it the highest lake in England. In 1992, the lake and its wetlands were designated as a National Nature Reserve. The estate was donated to the National Trust in 1947, which leases part of the site to the Field Studies Council who offer residential and non-residential field courses there. The area covered by the research was contained within the boundaries selected by C. A. Sinker for his study of the plant communities (Sinker, 1960). The major component are bog communities, woodlands on raised – bog peats and small fens, fen meadows and fen carrs. The four selected sampling sites used in this study were: Ha Mire, Tarn Woods, Tarn Fen and Spiggot Hill.

Ha Mire

Ha Mire is an area of flat sloping mire in the eastern part of the Tarn covered by calcareous marshes. This area has a large plantation (10 acres) which was established during the 19th century (Seaward and Pentecost, 2001). It is a small, fenced wood of mature and young trees circled by a high dry-stone wall, designed to protect the trees from grazing sheep and cattle. The vegetation of Ha Mire is dependent on calcareous groundwater and rainfalls. Most of the plantation is covered by grass (*Agrostis* spp., *Festuca* spp.), bracken (*Pteridium* spp.) and bramble (*Rubus* spp.). The woodland is dominated by sycamores (*Acer pseudpolatanus*), alders (*Alnus* spp.), larches (*Larix* spp.) and birches (*Betula* spp.).

Figure 1.7 Ha Mire plantation, view from the distance and inside (Photos taken by J. Bajnok and K. Boyce)

Tarn Woods

Tarn Woods is the largest plantation (40 acres) situated on the north shore of the tarn and surrounding the Field Centre. It is mixed deciduous and coniferous woodland (Sinker, 1960) including larch (*Larix* spp.), spruce *(Picea* spp.), ash (*Fraxinus excelsior*), rowan tree (*Sorbus aucuparia*) and beech (*Fagus* spp.). The ground cover is dominated by dog's mercury (*Mercurialis perennis*), large bryophytes, grasses (*Poa trivialis)*, goosegrass *(Gallium aparine) and* red campion *(Silene dioica*).

Figure 1.8 Tarn Woods (Photos taken by K. Boyce)

Tarn Fen

Tarn Fen lying to the north of Tarn Moss is woodland amongst bog/marsh on the edge of the tarn and contains a diversity of wetlands sustained by springs and overbank flooding from the stream. This area covers an area of 26 acres and consists of three parts: West, Middle and East Fen. The major components are bog communities dominated *by Eriophorum vaginatum* and *Calluna vulgaris*, *Molinia caerulea* vegetation and birch woodlands (*Betula pubescens*) with bushy willows (*Salix* spp.) and alders (*Alnus* spp.) on the raised bog peats. The rich fens and associated swamps contain some tall- herb vegetation (*Potentilla palustris, Angelica sylvestris, menthe acquatica, Carex rostrata, Urtica dioica and Acrocladium nodum*). Some areas of acidic peats are occupied by *Carex echinata* and *Sphagnum recurvum* (Sinker, 1960). It is only East Fen that has peats permanently kept wet with calcareous groundwater. Locally we can see some fallen trunks. The differences in vegetation structure in Tarn Fen reflect the differences in water levels. Almost the whole area shows substantial variations in water levels with the seasons and with changing weather. The variations are greatest near the stream inflows and around the margins of the fen.

Figure 1.9 Tarn Fen (Photos taken by K. Boyce)

Spiggot Hill

Spiggot Hill is located at the southwest junction of the Tarn's south and west shores. The vegetation is influenced by hydrological inputs, including streams that flood periodically. Several small peat pools are located to the northwest of Spiggot Hill. The area is covered by deciduous and coniferous woodland (Sinker, 1960) with larch (*Larix* spp.), birch (*Betula* spp.) and sycamore (*Acer pseudoplatanus*). The ground flora contains mainly grasses, dog's mercury and meadowsweet (*Filipendula ulmaria*).

Figure 1.10 Traps set in Spiggot Hill (Photos taken by K. Boyce)

Figure 1.11 Spiggot Hill in winter (Photos taken by K. Boyce)

To date only few studies have been conducted on *A. sylvaticus* infections by *T. gondii*. Serological analysis of similar species, the striped field mice (*Apodemus agrarius*) detected prevalence of 1.49% in Korea (Jeon *et al.* 2000) and 7.4% in the Czech Republic (Hejlicek *et al.* 1998). Higher prevalences were reported after testing wood mice with PCR methods, with prevalences ranging from 14.3% (Kijlstra *et al.* 2008) up to 40.78% (Thomasson *et al.* 2011).

In this study we set out to analyse prevalence of *T. gondii* in populations of wood mice at the family level in relation to *T. gondii* infection and to use direct genotyping methods to identify the parasite strains circulating in wildlife in UK. This research is fully described in Chapter 3 and Chapter 4.

1.6 Lung cancer

Lung cancer (pulmonary carcinoma) is a malignant lung tumour characterized by uncontrolled cell growth in tissues of the lungs and the most common cancer in both men and woman in the world (Parkin *et al.* 2010). There were an estimated 1.2 million new cases and 1.1 million deaths in 2000 (Parkin *et al.* 2010). The highest rates are in North America, Europe and East Asia, with over a third of new cases in 2012 in China (Spiro *et al.* 2010). In the European Union, there has been a decline in lung cancer mortality in men but the opposite situation was observed for women, in whom lung cancer rates increased from 9.0 to 11.4 cases per 100 000, an increase of 27% (Rodriguez *et al.* 2012). In England and Wales the increase was 4% (La Vecchia *et al.* 2010). More than 90% of adults with lung cancer are asymptomatic when they are diagnosed. There are two main types of lung cancer: Small cell lung cancer (SCLC) and Non-small lung cancer (NSCLC). SCLC often starts in the bronchi near the centre of the chest and spreads widely trough the body. NSCLC includes about 85-90% of all lung cancers (American Cancer Society, 2014). There are three main subtypes of NSCLC: squamous cell (epidermoid) carcinoma, adenocarcinoma and large cell carcinoma. Along with the two main types of lung cancer, other tumours can occur in the lungs: lung carcinoid tumours or cancer that has spread to the lungs. It has been observed that 10% of men and 20% of women that develop lung cancer are non-smokers (Sasco *et al.* 2004), so it has been suggested that the lung cancer in men can be considered a different disease from lung cancer in women (Sun *et al.* 2007).Women with lung cancer tend to be younger and have lower smoking histories than men, and are more likely to develop adenocarcinoma (Olak *et al.* 2004). Many observations suggest that hormones, especially oestrogens may be involved in lung carcinogenesis (Stabile *et al.* 2002). There are several risk factors that can increase the chance of getting lung cancer. Smoking is the leading risk factor. At least 80% of lung cancer deaths are thought to result from smoking with secondary smoking causing more than 7000 deaths from lung cancer each year (American Cancer Society, 2014). Exposure to radon, asbestos fibres, diesel exhausts, radioactive ores cadmium, and arsenic can increase risk of getting lung cancer (American Cancer Society, 2014). Polymorphism on chromosomes 5, 6 and 15 are known to affect the risk of getting lung cancer (Kern *et al.* 2008). Patients with lung cancer suffer from progressive shortness of breath, cough, and pneumonia, loss of voice and chest pain. The diagnosis is confirmed by biopsy which is usually performed by bronchoscopy. Common treatment includes surgery, chemotherapy and radiotherapy. In England between 2005 and 2009, the overall five year survival for lung cancer was less than 10% (Khemasuwan *et al.* 2015).

1.7 Chronic obstructive pulmonary disease (COPD)

Chronic obstructive pulmonary disease is the most common respiratory condition and is characterized by decline in lung function over time. The main symptoms include shortness of breath, cough and sputum production (Vestbo *et al.* 2013). Acute exacerbations are associated with a rapid decline in lung function and are the major causes of morbidity and mortality in COPD (Beasley *et al.* 2012). The inflammation in the small airways often intensifies with disease progression and is characterized by infiltration of inflammatory and immune cells, CD4+, T-cells, B-cells, macrophages and neutrophils (Di Stefano *et al.* 1996). Currently the prevalence is estimated to be around 10% in the population aged over 40 years (Gagnon *et al.* 2014). Smoking is the leading factor, but only 15%-20% of the smokers develop the disease (Fletcher *et al.* 1977). In England around 0.84 million people have been diagnosed with COPD (WHO, January 2015). In many studies the incidence of COPD was greater in men than in women (de Marco *et al.* 2007). Most cases of COPD are preventable through decreasing exposure to smoke and improving air quality. Inhaled bronchodilators are the primary medication used. There are several short acting β-agonists (salbutamol, terbutaline) or long acting β-agonists (salmeterol, formoterol). Corticosteroids are usually used in inhaled form or tablets and prevent acute exacerbations. Patients with severe exacerbations often use antibiotics to improve the outcomes.

Patients with cancer may have deficient cellular immunity and are potentially susceptible to *T. gondii* infections (Montoya *et al.* 2004). Not much is known about toxoplasmosis in this group of patients and few reports are available. Serological measurement of infection rates (Yuan *et al.* 2007) showed high prevalences in nasopharyngeal carcinoma (46.15%) and rectal cancer (63.64%) but lower rates in the other cancer groups, for example, pulmonary carcinoma (4.55%), breast cancer (9.53%), gastric carcinoma (10.00%), hepatocellular carcinoma (14.29%) and uterine cervix carcinoma (12.50%). These studies have little indication as to whether active infection is present, and they measure generic infection status rather than localised infection status in the cancer affected tissue. In this study specific DNA based and immunohistochemical detection systems were used to detect the presence of *T. gondii*, in lung biopsy samples taken from a well-characterised collection of patients with cancer.

The spread and clinical manifestations of an infection in human populations depends on a variety of factors, among them host genetics. As humans migrated throughout the world, populations encountered distinct pathogens, and natural selection increased the prevalence of alleles that are advantageous in the new ecosystems in both host and pathogens. In this study we have used new approaches to study any possible links between *T. gondii* and a variety of lung diseases including lung cancer. Similar approaches to the way we studied *T. gondii* prevalence in natural mice populations were used. A combination of molecular and histological methods was used to investigate the prevalence of *T. gondii* in populations of humans with lung cancer. Toxoplasmosis has been only recently recognized as significant cause of mortality at a population level for southern sea otters (Miller *et al.* 2008). Another study suggests that *T. gondii* has also a negative effect on the population dynamics of Hector's dolphins (Roe *et al.* 2013). These studies show that *T. gondii* can be important agent in the population decline of endangered species, and highlights the need for further research into the route of transmission. As most of the published data is based on studies performed in laboratory or domestic animals, there is lack of information on prevalences and genotypes of *T. gondii* circulating in wild life and human populations. Procedures applicable to and knowledge acquired from studies related to these animals can be extrapolated to wild life and humans. Deeper analyses and understanding of mechanisms of transmission in the sylvatic cycles and natural populations are necessary to better understand the possible links between individual strains and disease symptoms.

There is a need for development of generic approaches which can be used in studies of animals and humans. In this study we conducted a detailed analysis of the ecology of the wood mouse and investigated the prevalence, genotypes and transmission of *T. gondii* in this model system. In the second part, we selected a natural sub-population of humans (lung cancer patients) to further investigate these approaches. We analyzed collected sampled tissue using molecular methods and used them to investigate host genotype and spatial location in relation to *T. gondii* infection.

Objectives

The aim of our study was to determine the prevalence and genotypes of *T. gondii* in natural rodent populations and in humans with a variety of lung diseases (cancer, asthma and COPD). Due to the usual diet of the wood mouse as granivores, testing for the presence of *T. gondii* in the area free of cats provides an opportunity to examine transmission routes that do not involve bradyzoite ingestion, and to confirm the presence of vertical transmission. In addition, vertical transmission of infection in mammals often ends with foetal abortion and is causing significant economic losses in the UK livestock industry. As rodents often serve as reservoir of *T. gondii*, it is important to know the actual prevalence of this parasite and its potential risk for the farm animals and humans. Cancer patients may have a compromised immune system and are potentially susceptible to *T. gondii* infections. Not much data are available on genotypes present in human populations as most of the studies were based on serological testing. Genetic analysis of *T. gondii* infecting humans is important to understand epidemiology, transmission patterns, and mechanisms of the disease. However, the difficulty in obtaining *T. gondii* strains from humans is a limiting factor. In order to identify the parasite genotypes; we have chosen to do direct PCR, sequencing and genotyping of the samples. Human samples were also tested with histological and immunohistochemical methods. We discuss the efficiency of individual sampling and detection methods, involvement of possible genotype pathogeny and potential sources of infection.

Chapter 2:

Materials and Methods

2.1. Sampling of *Apodemus sylvaticus*

All samples were collected from surrounding area of Malham Tarn Research Field Centre as a part of an over 10 year ongoing study of parasites in this area. A total of 126 *A. sylvaticus* were collected over a two year period (October 2009- October 2011) using Longworth Traps and euthanased by anaesthetic inhalation, from four sites located within the boundaries of the Malham Tarn Nature Reserve, North Yorkshire, UK (Figure 2.1.) as described previously (Boyce *et al.* 2012; 2013). Collection points at these sites, labelled Tarn Woods (54^{o}06’03.3”N, 002^{o}09’44.9”W), Tarn Fen (54^{o}06’00.0”N, 002^{o}10’43.4”W), Ha Mire (54^{o}05’72.9”N, 002^{o}10’53.7”W) and Spiggot Hill (54^{o}05’72.9”N, 002^{o}10’43.1”W) were recorded using GPS position fixing (Datum: WGS84). Trapping was carried out over a period of four nights at each site. Four trap lines consisting of 15 traps were set up with a 3m distance between two traps and when possible were placed strategically alongside and resting on top of natural objects such as fallen tree trunks and branches within shrubs and grass tussocks. All appropriate permissions were obtained (Boyce *et al.* 2012; 2013) and ethical approval was granted by the University of Salford Research Ethics and Governance Committee (CST 12/36). Mice were examined for a range of parameters including sex, weight and length. Mice weighing less than 14 g were considered juveniles (Higgs and Nowell, 2000). The brains were dissected out, using sterile technique, and transferred into sterile tubes containing 400 µl of lysis buffer (0.1 M Tris pH 8.0, 0.2 M NaCl, 5 mM EDTA, 0.4% SDS) and stored at −20 °C until DNA extraction (Terry *et al.* 2001). Neither serological analysis for *T. gondii* infection nor isolation of viable parasites was possible due to sampling logistics.

Figure 2.1 Location of Malham Tarn within the North Yorkshire (www.ciaofamiglia.com/frfletcher/fletchlinks.htm, accessed 6/6/2016)

2.2. PCR detection of *Toxoplasma gondii* and RFLP genotyping

DNA was isolated, from *A. sylvaticus* brain tissue, using proteinase K lysis followed by phenol/chloroform extraction as previously described (Duncanson *et al.* 2001). Extracted DNA was tested by PCR for mammalian tubulin gene amplification to ensure the viability for subsequent PCR (Terry *et al.* 2001) and appropriate protocols to prevent cross contamination were followed (Williams *et* al. 2005; Hughes *et* al. 2006; Morley *et al.* 2008). Detection of *T. gondii* was carried out using nested PCR amplification of the SAG1 genes (Savva *et al.* 1990) as modified by Morley *et al.* (2005). Positive amplification was confirmed by nested PCR amplification with three other sets of *T. gondii* specific primers (SAG2, SAG3 and GRA6) as described by Su *et al.* (2006) and Shwab *et al.* (2013). All samples were tested as positive three times with each of the four markers before the mouse brain was considered positive for *T. gondii* infection. In addition to being used for parasite detection, these three genes (SAG2, SAG3 and GRA6) were used as RFLP markers for direct genotyping of PCR positive brain tissues as described (Su *et al.* 2006; Shwab *et al.* 2013). *Toxoplasma gondii* infections were also tested by nested PCR amplification of the repetitive and conserved gene B1 (Jones *et al.* 1999). *Toxoplasma gondii* DNA strains RH (Type I), SR (Type II) and C56 (Type III) were used as positive controls for diagnostic PCRs and genotyping and sterile water was used as a negative control. Negative controls were included at all stages to ensure that any contamination would be detectable All PCR reactions were performed using a Stratagene

ROBOCYCLER™ (La Jolla, California, USA). PCR products were run on agarose TBE gels containing GelRed and visualized on a SYNGENE G: Box Gel Documentation and Analysis System (Cambridge, UK). To ensure the reliability of the genotyping PCRs, the following procedures were followed for all three markers. Each sample was successfully amplified, independently, for a minimum of 3 times before the final genotype result was accepted.

2.2.1 Mammalian α-tubulin PCR

All 126 DNA samples were tested for mammalian tubulin to ensure the viability for PCR (Terry *et al.* 2001). Sheep DNA was used as a positive control. The PCR reaction mix contained 2.5 µl of HT PCR Buffer, 0 .5 µl of each primer (25pM), 0.25 µl dNTP mix (25mM each), 0.5 µl Taq DNA polymerase (5U/µl) and 19.75 µl of PCR water. Amplification was carried out using a Stratagene Robocycler as follows: an initial denaturation step of 5 min at 94 °C was followed by 40 cycles of PCR performed for 40 sec at 94 °C, 40 sec at 60 °C and 1 min 30 sec at 72 °C, with a final extension step of 10 min at 72 °C. The PCR product was run on a 1% TBE gel using 1 kb marker (Bioline). The following primers were used $TUB_{FOR:}$ $5^{/}$CGTGAGTGCCATCTCCATCCAT-$3^{/}$ and TUB_{REV}: $5^{/}$GCCCTCACCCACATACCAGTG-$3^{/.}$

2.2.2 Amplification of the B1 Gene

Nested primers targeting the B1 gene (Jones *et al.* 2000) included outer primers for the first round of amplification, consisting of forward primer $5^{/}$GGAACTGCATCCGTT CATGAG-$3^{/}$ and reverse primer $5^{/}$ TCTTTAAAGCGTTCGTGGTC-$3^{/}$ which generates a 193 bp product (nucleotides 694-887). The inner set of second round of primers consisting of forward primer $5^{/}$ TGCATAGGTTGCAGTCACTG-$3^{/}$ and reverse primer $5^{/}$GGCGACCAATCTGCGAATACACC-$3^{/}$ generates a 96 bp product after amplification by nPCR.

PCRs for the first round contained 10mM TRIS-HCl, pH 8.3, 50mM KCl Buffer, 2mM $MgCl_2$ (Bioline KCl buffer), 0.1µM each primer, 0.1mM each dNTP, 1.25 U Taq DNA polymerase and 2 µl of parasite DNA. Reactions were cycled 40 times with denaturation at 93˚C for 10 seconds followed by annealing at 57˚C for 10 seconds and finally an extension step at 72˚C for 30 seconds. A final extension step of 10 min at 72 °C was added (not used in the original protocol).

Second round reactions contained 1µ of the first-round product, 10mM Tris-HCl, pH 8.3, 50mM KCl, 3mM $MgCl_2$, 0.5 µl each primer, 0.1mM each dNTP and 1U Taq DNA polymerase. Nested PCRs were cycled 40 times using a denaturation step of 93˚C for 10 seconds, followed by annealing at 62˚C for 10 seconds and extension at 72˚C for 15 seconds. Negative control samples from first-round amplification and were included in the nested reactions. A final extension step of 10 min at 72 °C was added. Ten µl of amplification products were visualized under UV illumination after electrophoresis on 3.5% TBE gel using Hyperladder V (Bioline) as a marker.

2.2.3 SAG1 PCR

The presence of *T. gondii* was detected using a nested PCR amplification of the Surface Antigen Gene 1 (SAG1) and resolved by gel electrophoresis as previously described by Williams *et al.* (2005). Each sample was tested at two dilutions of DNA (1/5 and 1/10 dilution as the ratio of parasite to host DNA was unknown). Four oligonucleotide primers were used. Their sequences were based on the published sequences for the toxoplasma P30 gene; two of these (DS29 and DS38) were of the same sense. Oligo DS29 ($5^{/}$TTGCCGCGCCCACACTGAG-$3^{/}$) corresponds to nucleotides 405-424 of the P30 gene, oligo DS30 ($5^{/}$CGCGACACAAGCTGCGATAG-$3^{/}$) to nucleotides 1318-1299, oligo DS38 ($5^{/}$GACAGCCGGCGGTCATTCTC-$3^{/}$) to nucleotides 503-522 and oligo DS39 ($5^{/}$GCAACCAGTCAGCGTCGTCC-$3^{/}$) to nucleotides 1024-1005. When used in different combinations (DS29-DS30, DS29-DS-39, DS38-DS39, and DS30-DS38) these primers should result in the amplification of fragments of *T. gondii* of 914, 620, 522 and 815 bp respectively (Savva *et al.* 1990). The PCR contained 2.5 µl Bioline PCR Buffer (excluding $MgCl_2$), 2.5 µl β-mercaptoethanol (50mM), 1µl Bioline $MgCl_2$ (50mM), 2.5 µl of each primers DS29 and DS30 (10pM each), 0.25 µl each dNTP (25mM each), 0.5 µl Taq DNA polymerase (5U) and 12.25 µl PCR water. Amplification was carried out using a Stratagene Robocycler as follows: an initial denaturation step of 5min at 95 °C was followed by 40 cycles of PCR performed for 40 sec at 95 °C, 40 sec at 63 °C and 1min 10 sec at 72 °C, with a final extension step of 10 min at 72 °C. Second-round PCR was carried out using the same reaction and cycling conditions as the first round (combination of primers DS38 and DS39 was used). A volume of 2 µl of first-round product was added to act as a template. Amplification products (10 µl) were visualized by agarose gel electrophoresis on a 1.5% TBE agarose gel containing GelRed. Hyperladder 100 bp (Bioline) marker was used as size ladder. 522 bp products at the end of second round of amplification were considered PCR positive.

2.2.4 SAG2 PCR

SAG2 locus has two polymorphic sites at $3^{/}$ and $5^{/}$ ends for Type II and Type III (Howe *et al.* 1997), so amplification of this locus was performed separately. Two nested PCRs, one targeting the $3^{/}$ end of the SAG2 gene and the other targeting the $5^{/}$ end of SAG 2 gene were optimized as per the method of Fuentes *et al.* (2001). The nucleotide sequence of the outer and the inner set of primers targeting the $3^{/}$end and the $5^{/}$end of SAG2 gene are listed below. Amplification was carried out in a final volume of 20 µl containing 2.7 µl of KCl Bioline buffer, 0.32 µl of dNTP mix (25mM each), 1 µl of (10 pM) forward primer and reverse primer and 0.4 µl of 5 unit Taq DNA polymerase (Bioline). Two µl of DNA were used as a template. First round amplification cycle was initiated with 4 minute denaturation at 95˚C, followed by 20 cycles amplification. Each cycle consisted of incubations at 94˚C for 30 s, 55˚C for 1 min and 72˚C for 2 min. At the end the reactions were incubated at 72˚C for 10 min. The resulting amplifications products were diluted 1/10 in water and a second amplification of 35 cycles was performed, using 2 µl of the diluted product as template and inner primer sets as designated below. The annealing temperature for the second round was 60˚C.

For the nPCR targeting the $3^{/}$ end of SAG2 gene, detection of 222 bp products at the end of second round of amplification were considered PCR positive. For the nPCR targeting the $5^{/}$ end of SAG2 gene, detection of 242 bp products at the end of second round of amplification was considered PCR positive.

Table 2.1 Primer sequence of nPCR targeting the 5′ and 3′ end of SAG2 gene (Howe *et al.* 1997; Su *et al.* 2006):

	SAG2 5′ end	SAG2 3′ end
External markers	5′ GACCTCGAACAGGAACAC3′	5′TCTGTTCTCCGAAGTGACTCC3′
	5′GACCTCGAACAGGAACAC3′	5′TCAAAGCGTGCATTATCGC3′
Internal markers	5′GAAATGTTTCAGGTTGCTGC3′	5′ATTCTCATGCCTCCGCTTC3′
	5′GCAAGAGCGAACTTGAACAC3′	5′AACGTTTCACGAAGGCACC3′

Positive PCR reactions were further analysed by restriction enzyme digestion with the restriction enzymes *MboI* (5′-end products) and *HhaI* (3′-end products), 8.5 µl of PCR product, 1.5 µl of manufacturers recommended restriction enzyme buffer and 0.5 µl of enzyme (Fuentes *et al.* 2001). Reactions were incubated at 37 °C for a minimum of 2h or overnight. Products were visualized by gel electrophoresis on a 2.5% TBE agarose gel and Hyperladder II (Bioline) was used as a size marker.

2.2.5 *SAG3 PCR*

A nested PCR was used to detect the SAG3 gene. This method has been modified from the method described by Grigg *et al.* (2001) and Su *et al.* (2006). Amplification was carried out in a final volume of 50 µl containing 5 µl of 10× HT PCR buffer (HT Biotechnologies) (100mM Tris HCl with pH 9.0, 15mM $MgCl_2$, 500mM KCl, 1% TritonX-100, 0.1% (w/v) stabilizer), 0.5 µl of dNTP mix (100mM), 2.5 µl of (10 pM) forward primer (Fext:5′CAACTCTCACCATTCCACCC-3′) and reverse primer (Rext: 5′GCGCGTTGTTAGACAAGACA-3′) and 2.5 units Taq DNA polymerase (Bioline). DNA-free water made the final volume to 50 µl. Amplification was carried out using a Stratagene Robocycler as follows: an initial denaturation step of 5 min at 94 °C was followed by 35 cycles of PCR performed for 40 sec at 94 °C, 40 sec at 60 °C and 60 sec at 72 °C, with a final extension step of 10 min at 72 °C. Second-round PCR was carried out using the same reaction and cycling conditions as the first round with the exception of the primers which were $F_{int:}$ 5′TCTTGTCGGGTGTTCACTCA-3′ and R_{int}: 5′CACAAGGAGACCGAGAAGGA-3′. A volume of 2 µl of first-round product was diluted into reaction mix (as above) to give a final volume of 50 µl. Amplification products (10 µl) were visualized by agarose gel electrophoresis on a 1.5% TBE agarose gel containing Gelred and Hyperladder I size marker (Bioline).

Positive PCR (226bp) reactions were further analysed by restriction enzyme digestion with each of the enzymes *Nci*I and *AlwN*I, 13 µl of PCR product, 1.5 µl of buffer 4 (NEB) and 0.5 µl of enzyme. Reactions were incubated at 37 °C for a minimum of 2 h or overnight. Products were visualized by gel electrophoresis on a 2.5% agarose gel using Bioline Hyperladdder V size marker (3.5µl). *NciI* cuts both Type I and Type III to generate distinct fragment patterns (Grigg *et al.* 2001; Su *et al.* 2006) while Type II remains uncut. To distinguish Type II pattern from undigested PCR products, *Alwn*I (which cuts only Type II) was used to confirm that uncut PCR products were indeed Type II).

2.2.6 GRA6 PCR

GRA6 nested PCR was performed using published sequences: $F_{1:}$ $5^{/}$ATTTGTGTTTCCGAGCAGG-$3^{/}$, R_1: $5^{/}$GCACCTTCGCTTGTGGT-$3^{/}$ and $F_{2:}$ $5^{/}$TTCCGAGCAGGTGACC-$3^{/}$, R_2: $5^{/}$GCCGAAGAGTTGACATAG-$3^{/}$ (Su *et al.* 2006) and an optimized protocol The PCR reaction is carried out in a vol. of 25 µl containing 2.5µl of Bioline NH_4Cl buffer (excluding $MgCl_{2)}$, 1µl of 50 mM $MgCl_2$ buffer (Bioline), 0.32µl of the dNTP (25mM each), 1.5µl of the external forward primer and 1.5µl of reverse primer (10pM each) and 0.32 Taq DNA polymerase (5U, Bioline). DNA-free water made the final volume to 25µl. The reaction mixture is treated at 95˚C for 4 min, followed by 30 cycles of 94˚C for 30 sec, 55˚C for 1.5 min and 72˚C for 2 min. A final extension step of 10 min at 72 °C was added (not used in the original protocol). The resulting amplifications products were diluted 1/10 in water and a second amplification of 35 cycles was performed, using 1 µl of the diluted product as template in reaction mixture (as above) containing primers F_2 and R_2. The annealing temperature for the second round was 60˚C.

A 344 bp product at the end of the second round was considered to be positive. For RFLP typing, 5 µl of PCR products were treated with *MseI* in total volume of 20 µl at 37˚C for 1 h and then digested samples were resolved on a 2.5% TBE agarose gel to reveal restriction fragment banding patterns (Khan *et al.* 2005).

2.3. Microsatellites analysis and host genotyping

A total of 126 DNA samples from *A. sylvaticus* were prepared for microsatellite genotyping. Nine *A. sylvaticus* loci were amplified and genotyped using fluorescent dye-labelled DNA primers (Table 2.2.) and conditions previously described (Makova *et al.* 1998). One *Mus domesticus* microsatellite, which amplifies from *A. sylvaticus* was also used (MS19) The PCR reaction mixture (20 µl) consisted of 2 µl Bioline buffer (excluding $MgCl_2$), 0.6 µl Bioline $MgCl_2$ (50 mM), 2 µl of each primer (10 pM/µl), 0.2 µl dNTP mix (25 mM each), 0.2 µl *DNA* Taq Polymerase (5U) and 12 µl PCR water.

Thermal cycling conditions used were an initial 5 min denaturation at 95 °C followed by 35 cycles at 95 °C for 40 sec (denaturation), 30–45 s at the annealing temperature (Table 2.3.), 72 °C for 30–60 s (extension), concluding with a final 30 min extension at 72 °C (Makova *et al*. 1998,). Amplification products (10 µl) were visualized by agarose gel electrophoresis on a 1.5% TBE agarose gel containing GelRed.

Table 2.2 Primer sequences (5′ to 3′), allele ranges and fluorescent dyes (green–Hex, blue-FAM) used for 10 microsatellite loci in *A. sylvaticus* (Makova, 1998).

Locus (Dye)	**Primer sequences**	**Allele range (bp)**
GTTA1A	F: 5′ TTTGATGCCTTGACTTTGATTACC-3′	168-209
(HEX)	R: 5′AATGCCAGTGGTGATTTTATTTGG-3′	
CAA2A	F: 5′AATTTGCCCTTAAGTGAGGAAG-3′	94-118
(FAM)	R: 5′GCAGTGACCCAGGAGAAATTACC-3′	
GACAB3A	F: 5′AGGGGAACCTCACAAATATAGGAAA-3	361-622
(HEX)	R: 5′GGCTTCCAATTTTGAACTACAGAGC-3′	
GTTC4A	F: 5′GTAAATGGCTAGAAGGAGAGAAGGTTC-3′	128-154
(FAM)	R: 5′TTCCCTGGAAACTATTTGGTAAATCC-3′	
GCATD7S	F: 5′CTGGGTCTTCTGCACTGTTCTTTACC-3′	199-296
(FAM)	R: 5′GCAGATGCCCACCTTCTGTAACCAA-3′	
GTTD8S	F: 5′TCTGAACAGTGGTAGATAATTAGAGCTTA-3′	101-110
(FAM)	R: 5′GAAACCGTTTGGTAAGATACTACAAAA-3′	
GTTD9A	F: 5′CCCAAAATTGCCTTCCTGTCAC-3′	202-216
(HEX)	R: 5′GGTCAGGATAGGCTGCATAGAAAG-3′	
GACAE12	F: 5′GTTTTGTTTGGGTTTCTGAGACTGAA-3′	252-268
(HEX)	R: 5′ACTCGGCTCTTACTTGGTAATCTTCC-3′	
TNF	F: 5′AGGAAATGGGTTTCAGTTCTCAGG-3′	105-124
(FAM)	R: 5′GGTCCCCACCAGGATTCTGTG-3′	
MS19	F: 5′CTGATTTGAGCCTGTGCA-3′	80-226
(HEX)	R: 5′TGCTCAATAAATACAGAGCAAAGC-3′	

Table 2.3 Core repeats and PCR conditions for 10 microsatellite loci in *A. sylvaticus* (Makova *et al.* 1998).

Locus	Cloned repeat	Annealing temperature (°C)	Annealing time (s)	Extension time (s)
CAA2A	$(CA)_{21}$	58	30	45
GTTA1A	$(GTT)_{7}$ GAA (GTT)	55	45	60
GACAB3A	$(GATA)_{2}$ $(GACA)_{5}$ (GATA)	55	45	60
GTTC4A	(GTT) T3 $(GTT)_{5}$ T_{3} $(GTT)_{2}$	55	45	60
GCATD7S	$(CA)_{6}$ $(GCAT)_{3}$ $(GCAC)_{3}$	60	30	45
GTTD8S	$(GTT)_{9}$	55	45	60
GTTD9A	$(GTT)_{14}$	60	30	45
GACAE12A	$(GACA)_{6}$ $(CA)_{6}$ (GA)	60	30	45
TNF	(CA) $(CA)_{17}$	60	30	45
MS19	NA	58	40	60

PCR products were then mixed with formamide and LIZ™ standard marker and genotyped on the GENOTYPER (Applied Biosystems 3130 Genetic Analyser) according to the manufacturer to gain allele sizes relative to an internal size standard. Each sample for each locus was scored with Peak Scanner ™ v1.0 to identify peaks and fragment sizes for application-specific capillary electrophoresis assays (Dodd *et* al. 2014). The program STRUCTURE *2.3.3* (Pritchard *et al.* 2000; Evanno *et al.* 2005) was used to analyse multi-locus microsatellite genotype data to investigate population structure. The parameters for the burnin period and the Markov Chain Monte Carlo (MCMC) were both set to 100 000 replicates as recommended. The ancestry model was set to the admixture model as the mice were not expected to conform to the requirements for a fully discrete population. The allele frequency model was set to independent model as recommended (Pritchard *et al.* 2000). The sampling location was not used as prior information in the analysis. Replicate simulations were run for models with increasing numbers of assumed genetic clusters (K = 1-10) and average over 3 runs. The most likely K was inferred from graphical representations of K values plotted against the estimated log posterior probability following the recommendations (Evanno *et al.* 2005). Individuals were placed in clusters based on a proportion of greater than 50% of membership (q) being assigned to the cluster. In some cases, the program STRUCTURE cannot detect subgroups with weaker probabilities when all data is included. It is common practice (e.g. Gelanew *et al.* 2010) to rerun each initial group separately through STRUCTURE to investigate substructuring. The program MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004) was used to check for the presence of microsatellite null alleles and scoring errors.

The program COLONY v2.0.5.0 (Wang *et al.* 2004) was used to estimate the full- and half-sib relationships of mice. COLONY is a computer program implementing a maximum likelihood method to assign/infer parentage and sibship among individuals using their multi-locus genotypes. COLONY can be used, among others programmes, in estimating full- and half-sib relationships, assigning parentage, inferring mating systems (polygamous/monogamous,) and reproductive skew in both diploid and haplo-diploid species. It can also be used to simulate genotype data with a particular sibship and parentage structure. The new project wizard saved the data with a specific file name in the project folder (adding column headers if necessary), and assembled all of the data and parameters into a single file named “Colony2.dat” which is saved in the project folder on completion of the data input process. Empirical Data Analysis was selected in the first step. In most cases, the default values of the parameters were used. Longer runs consider more configurations in the simulated annealing process searching for the best assignment with the maximum likelihood, and thus are more likely to find the maximum likelihood configuration, but take more time to do so. Four methods (full likelihood, FL; pairwise-likelihood score, PLS; FL and PLS combined method, FPLS; pure pairwise likelihood, PPL) for sibship and parentage assignments are implemented in this program. We selected the FL method which is the most accurate one as verified by simulated and empirical data analyses (Wang *et al.* 2004). For the same dataset and parameters of a project, multiple runs can be conducted so that the best configuration with the maximum likelihood is more likely to be found and the uncertainties of the estimates are more reliably assessed.

The program BOTTLENECK 1.2 (Piry *et al.* 1999) was used to test a possible recent population bottleneck. This program generates the expected H_{EQ} (expected heterozygosity under mutation-drift equilibrium) from the number of alleles at a locus and the sample size using the SMM (stepwise mutation model), TPM (two-phase model) and IAM (infinite allele model) models. The H_{EQ} values are then averaged across loci and compared with the observed level of heterozygosity. The SMM counts new alleles along a stretch of DNA with respects to the addition or subtraction of particular subsets of DNA motifs, while the IAM considers any point mutation along a stretch of DNA within a locus to constitute a new allele.

The SMM and TPM are most appropriate for microsatellite data (Luikart *et al.* 1998). Five input data file formats (all are text files) are acceptable and automatically recognized by BOTTLENECK. One is GENEPOP (Raymond *et al.* 1995), the second is GENETIX (Belkhir *et al.* 1996) and the other three are described in the help file of the program. For our study the GENEPOP format was used. To achieve high statistical power at least 10 polymorphic loci (microsatellites or alloenzymes) and sampling at least 30 individuals is recommended. For fewer than 20 loci, the Wilcoxon's test is the most powerful and appropriate while the standardized differences test is more suitable when using 20 or more loci (Cornuet *et al.* 1996). This test relies on the assumption of allele neutrality and mutation-drift equilibrium. If a population has been subjected to a recent bottleneck, the mutation-drift equilibrium is temporarily disrupted and H_e will be significantly greater than H_{EQ} calculated from the number of alleles (Nei *et al.* 1975). The Wilcoxon test allows the detection of the heterozygosity excess due to the faster loss of low frequency. In non-bottleneck population about 50% of the loci are expected to have a slight heterozygote excess and 50% slight heterozygote deficiency due to genetic drift or sampling error (Luikart *et al*. 1998). The number of iterations influences the precision of the H_{EQ} estimates. A minimum of 1000 iterations is recommended. The distribution of the heterozygosity obtained through the simulation also enables the computation of *P* value for measured heterozygosity.

The NEESTIMATOR V2.01 software was used to calculate genetic effective population sizes (*N*e) from the allele frequency data using multi-locus diploid genotypes from population samples (Nomura *et al.* 2008). We used the genotypic data in GENEPOP format (FSTAT is also accepted). NeEstimator V2.01 included three single-sample estimators (updated versions of the linkage disequilibrium and heterozygote-excess methods, and a new method based on molecular coancestry), as well as the two-sample (moment-based temporal) method. One or multiple methods for calculating N_e can be performed simultaneously generally performed on a single data input file. For the single-sample method (linkage disequilibrium, heterozygote-excess and molecular coancestry), each sample is treated as separate population.

The analysis of the population history was run on the computer program DIYABC 2.1.0 (Cornuet *et al.* 2014), which uses the approximate Bayesian computation (ABC). This Monte-Carlo algorithm is used to make model based interference on complex evolutionary scenarios that act on natural populations. The evaluation of evolutionary scenarios includes the comparison of competing scenarios and the quantification of their relative support, and the estimation of parameter posterior distribution under a given scenario. DIYABC allows considering complex population histories involving any combination of population divergences, admixtures and population size changes, with population samples potentially collected at different times. The main limitation of the current version is the absence of migration among populations after they have diverged.

Possible migrating individuals were detected with GENECLASS2 (Piry *et al.* 2004). Migrant detection needs a single data file that includes both the populations for which migrants will be searched and the potential source populations for migrants.

Statistical analysis

Statistical analysis was conducted using GRAPHPAD. *P*-values for expected and observed frequencies were calculated using the Chi-Squared test or using Chi-Squared analysis of contingency tables (r x c). Pairwise analyses were conducted using 2 x 2 contingency tables with P values calculated using the Fishers exact test (FET).

2.4. Processing and analysis of human lung tissue samples

The tissue samples were collected from 88 patients (provided by Helen Carlin as a part of her PhD study) with lung cancer at the Wythenshawe Hospital, Northwest Lung Centre in Manchester. For each sample, data was available on age, gender, lung conditions (COPD, asthma) if present and on smoking status of individuals. DNA from all 88 patients was used for molecular testing and only 76 tissue samples were of suitable quality to be used for immunohistochemical staining. The study received ethical approval from the South Manchester Local Research Ethics Committee (03/SM/396) and the University of Salford (CST 12/37). The lungs tissue was washed in sterile Phosphate Buffered Saline (PBS) and fixed in 10% formalin. After the fixation the tissue was embedded in paraffin. This was done on the Leica TP1020 automatic tissue processor. First, the tissue was dehydrated in range of alcohols, 50%, 70%, 90% and absolute ethanol for 1 hour each. After the dehydration step the tissue was left in two buckets of Histoclear, 1 hour each. In the final step the tissue was transferred into a paraffin bath (58˚C), 2 x 1.5h. Tissue blocks can be stored at room temperature for years. Tissues were sectioned using a microtome with the dial set to cut 5μm sections.

2.4.1 Haematoxylin and Eosin staining

Haematoxylin is a basic dye and has an affinity for the nucleic acids of the cell nucleus. Eosin is an acidic dye with an affinity for cytoplasmic components of the cells. The slides were stained using an automated stainer. First, they were dewaxed in Histoclear (2 x 5min) followed by rehydration in ethanol, 100% (5 minutes), 90% (3 minutes), 75% (2minutes), 50% (1 minute) and rinsed in distilled water for 5 minutes. The next step was staining with haematoxylin (6 minutes). Stained slides were rinsed under the running tap water for 5 minutes, decolorized in acid alcohol with one dip and rinsed again under the tap water. Tissue was counterstained with eosin (15 seconds). In final step, stained slides were dehydrated with alcohols, 50 % (1 minute), 75 % (2 minutes), 95 % (4 minutes), and 100 % (5 minutes), cleared in Histoclear (2 x 5 minutes) and mounted with cover slips in DPX.

2.4.2 Immunohistochemistry (IHC) for *T. gondii* in human lung sections

Immunohistochemistry was performed on paraffin embedded tissue using commercial polyclonal antibodies produced in rabbits (Thermo Fisher Scientific). The tissue sections were cut at 5 µm and mounted on positively charged glass slides. Slides were dewaxed in Histoclear (2 x 5min), rehydrated in alcohols (ethanol), 100% (5 minutes), 90% (3 minutes), 75% (2 minutes), and 50% (1 minute). The rack was finally rinsed in tap water to remove the ethanol for 5 minutes. Antigen retrieval was performed in 1% trypsin/calcium chloride (pH 7.8) at 37˚C, 30 minutes in a humidified chamber as recommended by Thermo Fisher. After the incubation the sections were left to cool down at room temperature for 10 minutes before they were washed in PBS Tween 20 for 2 x 2 minutes. Endogenous peroxidase activity was blocked by incubating slides in 0.3% hydrogen peroxide for 30 minutes at room temperature followed by washing in TBS. Nonspecific antibody was blocked using normal goat serum (Vectastain ABC Systems, Vector Laboratories, UK) for 30 minutes at room temperature and followed by incubation in diluted (1/100) polyclonal rabbit anti- *T. gondii* antibodies (Thermo Fisher Scientific, Catalogue number PA1-38789, Rockford, IL, USA) for 1 hour at room temperature. After the incubation, the slides were washed in TBS Tween for 3 x 3 minutes and incubated in Vector biotinylated anti-goat secondary antibody (Vectastain ABC Systems, Vector Laboratories, UK) for 30 minutes at room temperature followed by another washing in TBS Tween (3 x 3 minutes). The resulting complex was visualized using 3-3'-diaminobenzidine (DAB) for maximum of 10 minutes (visualised under the light microscope and the reaction stopped with dH_2O). Before the DAB visualization the slides were incubated in ABC-Px mix (Elite Px) mix for 30 minutes. Sections were then washed under the running water for 5 minutes, counterstained with haematoxylin for 45 seconds, washed under the water for another 5 minutes, dehydrated with alcohols, 50 % (1 minute), 75 % (2 minutes), 95% (4 minutes), and 100 % (5 minutes), cleared in Histoclear (2 x 5 minutes) and mounted with cover slips in DPX. Two negative antibody controls were used, lung sections from *T. gondii* negative wood mouse and human lung sections omitted of primary antibodies. Immunostained slides were visualised using a light microscope (Leica) at x400 magnification, slides were also assessed using quantitative criteria (ImageJ) to provide a percentage score which described the degree of infection. For each patient (slide) three randomly selected x400 magnification microscope fields were photographed. After calculation of the proportional staining area to whole fields with ImageJ (as percentage of stained pixels), the mean percentage of pixels in the stained areas was calculated for each slide. According to the final mean (intensity), the patients were divided into three grades. Grade 1 had intensity below 10%, grade 2 between 10-20% and grade 3 contained all the patients with staining intensity greater than 20%. Data was analysed by logistic regression. All analyses were undertaken using R (R Core Team, 2013).

2.5 PCR detection and genotypes of *T. gondii* in human samples

DNA from 88 patients was isolated from small blocks of frozen tissue, using proteinase K lysis followed by phenol/chloroform extraction as previously described (Duncanson *et al.* 2001). Extracted DNA was tested for mammalian tubulin gene to ensure sample viability for PCR (Terry *et al.* 2001). The presence of the parasite was tested with standard *T. gondii* markers SAG1, SAG2 ($3^{/}$ and $5^{/}$ were tested separately), SAG3 and B1. The protocols were the same as those used for testing *A. sylvaticus* and are fully described at the beginning of this chapter in section 2.2. Each sample was tested three times at its original DNA concentration. The sample was accepted as positive if the presence of the parasite DNA was confirmed in all three reactions. Additional strain typing with commonly used markers (Table 2.4) was performed (Su *et al.* 2006; Shwab *et al.* 2013). Samples were accepted as positive if at least one reaction gave positive results with listed markers. If samples were negative in all three reactions they were considered to be negative with that marker. All markers were tested on human genomic DNA (Promega, Madison, WI, USA, product code G304A) as negative control. Type II strain from Slovakia and RH (Type I) strain were used as a positive controls. Two negative controls were used, sterile water and DNA from wood mouse negative for *T. gondii*.

Selected tissue samples from two patients were sent to Salford Analytical Services (SAS) at the University of Salford for analysis with a Scanning electron microscope (SEM) and a Transmission electron microscope (TEM).

Table 2.4 *Toxoplasma gondii* PCR markers and RFLP conditions. The reaction mixture is treated at 95˚C for 4 min, followed by 35 cycles of 94˚C for 30 sec, 60˚C for 1 min and 72˚C for 1.5 min (for marker Apico, the annealing temperature is at 58˚C instead of 60˚C). For RFLP typing, 6-10 µl of nested PCR products were treated with restriction enzymes and digested samples were resolved in an agarose gel to reveal DNA banding patterns (Su *et al.* 2006), Grigg *et al.* 2001, Shwab *et al.* 2013)). Buffers for the restriction enzymes were provided by supplier (New England BioLabs). F= forward primers, R= reverse primers.

Markers	External PCR markers	Internal PCR markers	Size (bp)	Restriction enzymes, incubation temperature and time
SAG1	F: GTTCTAACCACGCACCCTGAG R: AAGAGTGGGAGGCTCTGTGA	F: CAATGTGCACCTGTAGGAAGC R: GTGGTTCTCCGTCGGTGTGAG	390	*Sau96I+HaeII* (double digest).37˚C 1 h. 2.5% gel
alt. SAG2	F: GGAACGCGAACAATGAGTTT R: GCACTGTTGTCCAGGGTTTT	F: ACCCATCTGCGAAGAAAACG R: ATTTCGACCAGCGGGAGCAC	546	*HinfI+TaqI*, 37˚C, 30min, 65˚C 30 min. 2.5% gel.
BTUB	F: TCCAAAATGAGAGAAATCGT R: AAATTGAAATGACGGAAGAA	F: GAGGTCATCTCGGACGAACA R: TTGTAGGAACACCCGGACGC	411	*BsiEI+TaqI* (double digest), , 60˚C 1 h. 2.5% gel
C22-8	F: TGATGCATCCATGCGTTTAT R: CCTCCACTTCTTCGGTCTCA	F: TCTCTCTACGTGGACGCC R:AGGTGCTTGGATATTCGC	521	*BsmAI+MboII* (double digest), 37˚C 30 min, 55˚C 30min. 2.5% gel.
C29-2	F: ACCCACTGAGCGAAAAGAAA R: AGGGTCTCTTGCGCATACAT	F: AGTTCTGCAGAGTGTCGC R:TGTCTAGGAAAGAGGCGC	446	*HpyCH4IV+RsaI* (double digest), 37˚C 1 h. 2.5% gel.
L358	F: TCTCTCGACTTCGCCTCTTC R: GCAATTTCCTCGAAGACAGG	F: AGGAGGCGTAGCGCAAGT R: CCCTCTGGCTGCAGTGCT	418	*HaeIII+NlaIII* (double digest), 37˚C 1 h. 2.5% gel
PK1	F: GAAAGCTGTCCACCCTGAAA R: AGAAAGCTCCGTGCAGTGAT	F: CGCAAAGGGAGACAATCAGT R: TCATCGCTGAATCTCATTGC	903	*AvaI+RsaI* (double digest), 37˚C 1 h. 2.5% gel.
Apico	F: TGGTTTTAACCCTAGATTGTG R: AAACGGAATTAATGAGATTTGAA	F: GCAAATTCTTGAATTCTCAGTT R: GGGATTCGAACCCTTGATA	640	*AflII+DdeI* (double digest), 37˚C 1 h. 3% gel.

CHAPTER 3:

Microsatellite-based genetic diversity and population structure of *Apodemus sylvaticus* in Malham Tarn Nature Reserve

3.1. Introduction

Mice from the genus *Apodemus* are the most common rodents of the Palaearctic region (Corbet e*t al.* 1978) and have shown diverse adaptations throughout their wide distribution range. Wood mice (*Apodemus sylvaticus*) and yellow-necked mice (*Apodemus flavicollis*) have been used in several microsatellite studies (Makova *et al.* 1998, Harr *et al.* 2000, Sommer *et al.* 2010) but nobody has used this approach to investigate host – parasite interactions. In our study we applied combined analysis of the host genetic markers and data based on the population structure in relation to parasite prevalence and genotype. Theoretical and analytical tools of population genetics have been widely applied for addressing various questions in the fields of ecological genetics and conservation biology. By looking at the variation of microsatellite markers in populations, inferences can be made about population structures and differences, genetic drift, genetic bottlenecks and even the date of a last common ancestor. Microsatellites are very good markers in genetic studies of populations because they have a high heterozygosity and so can be used to detect sudden changes in populations, effects of population fragmentation and interaction of different populations and they are useful in identification of new and incipient populations (Varilo *et al.* 2003). Microsatellites are short tandem repeats (STRs) or variable number tandem repeats (VNTR), which are made up of short repetitive sequences and they tend to occur in non-coding DNA. The repeated sequence is often simple, consisting of two, three or four nucleotides (di-, tri-, and tetranucleotide repeats respectively), and can be repeated 3 to 100 times, with the longer loci generally having more alleles due to the greater potential for slippage (Brinkmann *et al.* 1998). The repeat units are 1-13 nucleotides in length and are usually repeated 5-20 times. Dinucleotide repeats are the most frequently used loci, because they are the most common arrangement of microsatellite repeats (Li *et al.* 2002). Their density varies widely with species, peaking around one locus per 5kb. The DNA from the animal kingdom is very rich in CA repeats, with a ubiquitous chromosomal distribution in humans and mice (Dib *et al.* 1996, Dietrich *et al.* 1996). Trinucleotide repeats are largely studied in connection with some human diseases and cancers (Sutherlands, 1992) and are often found within exons. Tetranucleotide repeats are often GATA/GACA repeats, which occur in many higher organisms though their density varies with species. In diploid organisms each individual animal will have two copies of any particular microsatellite segment (Berger-Wolf *et al.* 2007). Over time, as animals in a population breed, they will recombine their microsatellites during sexual reproduction and the population will maintain a variety of microsatellites that is characteristic for that population and distinct from other populations which do not interbreed. Alleles at a specific location (locus) can differ in the number of repeats. The number of repeats at a locus can be determined by carrying out a PCR using primers that anneal either side of the STR, and then examining the size of the product by agarose, polyacrylamide gel electrophoresis or capillary electrophoresis. If you flank a microsatellite with fluorescent PCR primers then amplification will give a pair of fluorescent allelic products which will vary in size according to their repeat length (Jouquand *et al.* 1999). A population might possess alleles which vary in size like this.

The variability of microsatellites is due to a higher rate of mutation compared to other neutral regions of DNA. These high rates of mutation can be explained most frequently by slipped strand mispairing (slippage) during DNA replication on a single DNA strand. The primary mutational mechanism is believed to be replication slippage (Ellegren *et al.* 2000): during replication of microsatellite region the strands may be displaced and then realign incorrectly, leading to insertion or deletion of a number of repeat units. Mutation may also occur during recombination during meiosis. Interruption of microsatellites, perhaps due to mutation, can result in reduced polymorphism. Variation in flanking regions is also important. First point mutation in primer annealing sites may cause the occurrence of null alleles. Frequencies up to 15% have been reported, and they are found in up to 25% of loci (Callen *et al.* 1993). Null alleles may affect the estimation of population differentiation, for instance, by reducing the genetic diversity within populations. Null alleles may be detected in population studies by careful testing against frequencies that are expected under the Hardy-Weinberg equilibrium, provided that heterozygote deficiencies have no other origin, such as the mating system (Viard *et al.* 1996).

Microsatellites are considered to be neutral markers because unlike other parts of the genome, they do not code for proteins and thus we can assume they are not under selection. Microsatellites developed for particular species can often be applied to closely related species, but the percentage of loci that successfully amplify may decrease with increasing genetic distance (Abdul-Muneer *et al.* 2014).

There is no up to date data on wood mouse populations in pastoral areas of northern and western England, Scotland or Wales. A study was carried out in Northern Ireland in a habitat, where there was 1.9 farms and 9.2 km of field boundaries per km^2, and they found a density of *A. sylvaticus* of 3.0 per ha in summer and 2.5 per ha in winter, with 99% of these in field boundaries and only 1% in buildings (Montgomery and Dowie 1993). A study on sand dunes in Scotland suggested that densities range from less than 0.5 per ha in spring to around 12 per ha in the autumn (Gorman and Zubaid, 1993). In urban areas, densities can be very high in isolated habitat patches due to restricted dispersal (Dickman and Doncaster 1987), although suitable habitat patches are scattered and overall densities in urban areas are much lower. Montgomery (1989) suggested that the average April-June density for wood mice in mixed deciduous woodland is about 7 per ha. In arable areas of Britain, seasonal variation in density is from 0.5 per ha in the summer to 17.5 per ha in winter, with winter peaks as low as 8.4 per ha (Green 1979; Wolton 1985; Attuquayefio, Gorman & Wolton 1986; Wilson, Montgomery & Elwood 1993; Tew & Macdonald 1993; Tew 1994). In mixed farmland on the Moray coast, Grampian, wood mouse numbers varied from 2 to 30 per ha (M.L. Gorman pers. comm.). These low spring densities in arable fields may simply reflect the large size of fields. A total pre-breeding population is estimated to be about 38,000,000; 19,500,000 in England, 15,000,000 in Scotland and 3,500,000 in Wales. The autumn population would be about 114,000,000 (Harris *et al.* 1995). In this study we aim to investigate the structure of natural population of the wood mice (*A. sylvaticus*) in Malham Tarn area using genetic analysis and statistical analysis.

Objectives

1) Microsatellite genotyping of 126 *Apodemus sylvaticus* DNA samples.

2) Analysis of genotypes to determine population structure and family units.

3.2. Materials and Methods

A total of 126 DNA samples from *A. sylvaticus* were prepared for microsatellite genotyping. Nine *A. sylvaticus* loci were amplified and genotyped using primers and conditions previously described (Makova *et al.* 1998). One *Mus domesticus* microsatellite, which amplifies from *A. sylvaticus* was also used (MS19). PCR products were then mixed with formamide and LIZTM standard marker and genotyped on the Applied Biosystems 3130 Genetic Analyser according to the manufacturer to gain allele sizes relative to an internal size standard (full protocol is described in Chapter 2). Each sample for each locus was scored with PEAK SCANNER TM v1.0 to identify peaks and fragment sizes for application-specific capillary electrophoresis assays (Dodd *et al.* 2014). Quantification of allele frequencies and analyses of linkage disequilibrium were conducted using GENEPOP 3.4 (Raymond and Rousset, 1995) and GEMALEX 6.502 (Peakall *et al.* 2012), a cross-platform package for population genetic analyses that runs within Microsoft Excel. The program STRUCTURE 2.3.3 (Pritchard *et al.* 2000; Evanno *et al.* 2005) was used to analyse multi-locus microsatellite genotype data to investigate population structure. The program MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004) was used to check for the absence of microsatellite null alleles and scoring errors. The program BOTTLENECK (Piry *et al.* 1999) was used to determine if the population displayed a genetic signature of having gone through a bottleneck during the study period. Program COLONY v2.0.5.0 (Wang, 2004) was used to estimate the full- and half-sib relationships of mice. The analysis of the population history was run on the computer program DIYABC 2.1.0 (Cornuet *et al.* 2014), which uses approximate Bayesian computation (ABC). Contemporary effective population size was estimated with NEESTIMATOR V2.01 (Do *et al.* 2014) using linkage disequilibrium method. Possible migrating individuals were detected with GENECLASS2 (Piry *et al.* 2004).

3.3. Results

The aim of this study was to identify and optimise a sufficiently large set of microsatellites to use with wood mice to investigate population genetics and family units. In total, 126 wood mice *A. sylvaticus* were genotyped using 10 polymorphic microsatellite loci identified by Makova (1998). Examples of successful PCR microsatellite amplification are presented in Figure 3.1. All tested loci showed significantly higher level of polymorphism, with a maximum of 33 alleles present with MS19 and a minimum of 13 alleles for locus 8S. The average number of alleles per locus was 23.6 (SD 6.47) which is significantly higher than reported in the previous study. The allelic richness as well as observed and expected heterozygosity for each locus is shown in Table 3.1. Mean expected heterozygosity (H_e) was high, averaging from 0.359 (8S) to 0.927 (MS19), but six loci showed heterozygote deficiency. Significant departures (P<0.001) from Hardy –Weinberg equilibrium were observed on all ten loci when genotypes from all sites were pooled. Since most of the loci showed an excess of homozygotes, these deviations could be linked to family structure (Wahlund effect). No locus pairs were consistently in linkage disequilibrium and the program MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004) found no evidence of null alleles so all loci were considered independent for further analyses.

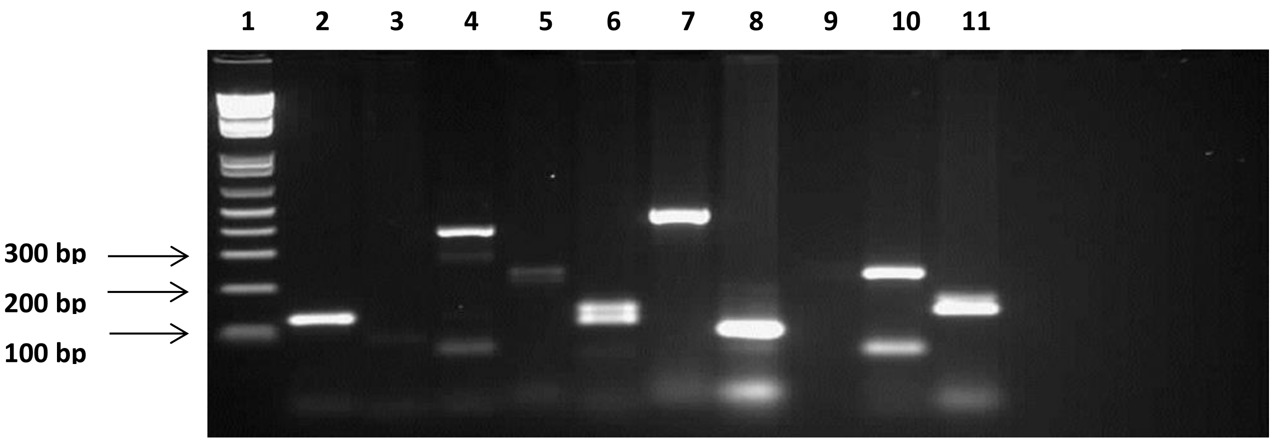


Figure 3.1 Example of a microsatellite PCR amplification of *A. sylvaticus* loci from sample number 12: Lane 1 = 1kb marker (Bioline), Lane 2 = 1A, Lane 3 = 2A, Lane 4 = 3A, Lane 5 = 4A, Lane 6 = 7S, Lane 7 = 8S, Lane 8 = 9A, Lane 9= 12A, Lane 10 = TNF, Lane 11 = MS19

Table 3.1 Microsatellite loci diversity for *Apodemus sylvaticus* sampled from Malham Tarn Nature Reserve. Microsatellites and expected allele range were originally defined by Makova *et al.* (1998).

Locus	***n***	***Na***	H_o	H_e	F_{ST}	**Allele range**
8S	122	13	0.364	0.359	0.071	101 – 110
TNF	118	28	0.924	0.919	0.043	105 – 124
4A	105	20	0.886	0.921	0.040	128 – 154
3A	107	27	0.935	0.903	0.039	361 – 622
1A	115	18	0.774	0.741	0.033	168 – 209
2A	116	25	0.491	0.795	0.241	94 – 118
7S	117	21	0.727	0.879	0.039	199 – 296
9A	111	19	0.478	0.745	0.047	202 – 216
12A	99	31	0.755	0.911	0.031	252 – 268
MS19	120	33	0.793	0.927	0.091	82-226
Mean	113	23.6	0.713	0.81	0.068	NA

n = number of individuals, out of 126, successfully genotyped, *Na* = number of detected alleles, H_e = frequency of expected heterozygotes, H_o= observed heterozygotes, F_{ST} = fixation index

The Bayesian analysis using STRUCTURE indicated the presence of multiple unique evolutionary clusters (R1, R2, B and G) of wood mice in Malham Tarn Nature Reserve. The highest probability of the data Pr (X/*K*), was calculated using STRUCTURE for *K*= 4 (Table 3.2, Figures 3.2. and 3.3.). The sampling location was not used as prior information in the analysis so as not to bias any genetic interpretation. As the generation of groups by STRUCTURE is based on probability of an individual belonging to a group, sometimes groups with weaker probabilities can be masked during the first analysis. It is common practice (e.g. Gelanew *et al.* 2010) to re-examine genetically defined STRUCTURE groups by re-running each group (i.e. R1, R2, G and B) separately through STRUCTURE. Following this more detailed analysis, STRUCTURE was unable to divide the individuals from all four populations into subpopulations so they were considered as individual populations. These genetic populations, identified by STRUCTURE, were examined to see if there was any association between microsatellite derived populations and the trapping location of the mice. Interestingly, the evidence of admixture showed that most of the sampling sites did not consist of one population, but that each location contains individuals from two or four generic clusters with different proportions (Figure 3.5).

Table 3.2 Inference of genetic clusters (*K)* from *A. sylvaticus*. The log-likehood values based on the STRUCTURE algorithm for each number of clusters are shown. The change in the value for *K*=4 is suggesting that most likely four evolutionary clusters are present.

K	1	2	3	4	5	6	7	8
-Log-likelihood	-5207	-5088	-5065	-5055	-5092	-5126	-5163	-5194

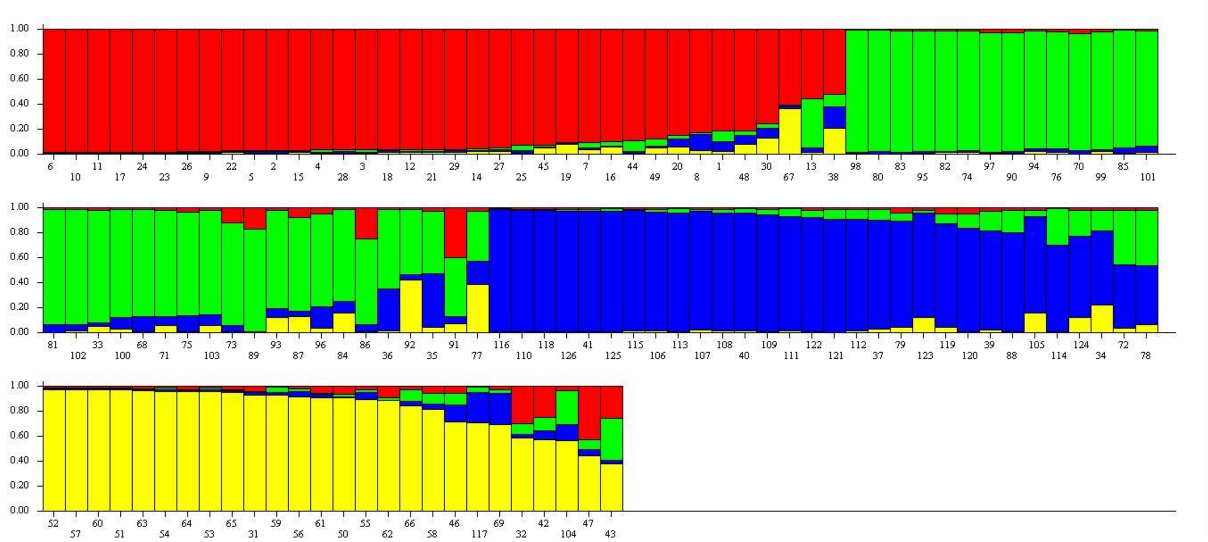


Figure 3.2 Four evolutionary clusters (*K*=4) inferred from STRUCTURE software (Pritchard *et al.* 2000; Evanno *et al.* 2005) analysis of 126 mice from four sampling sites. Each individual is represented by a vertical line broken into segments of different colours, with lengths proportional to its membership to each of the clusters.

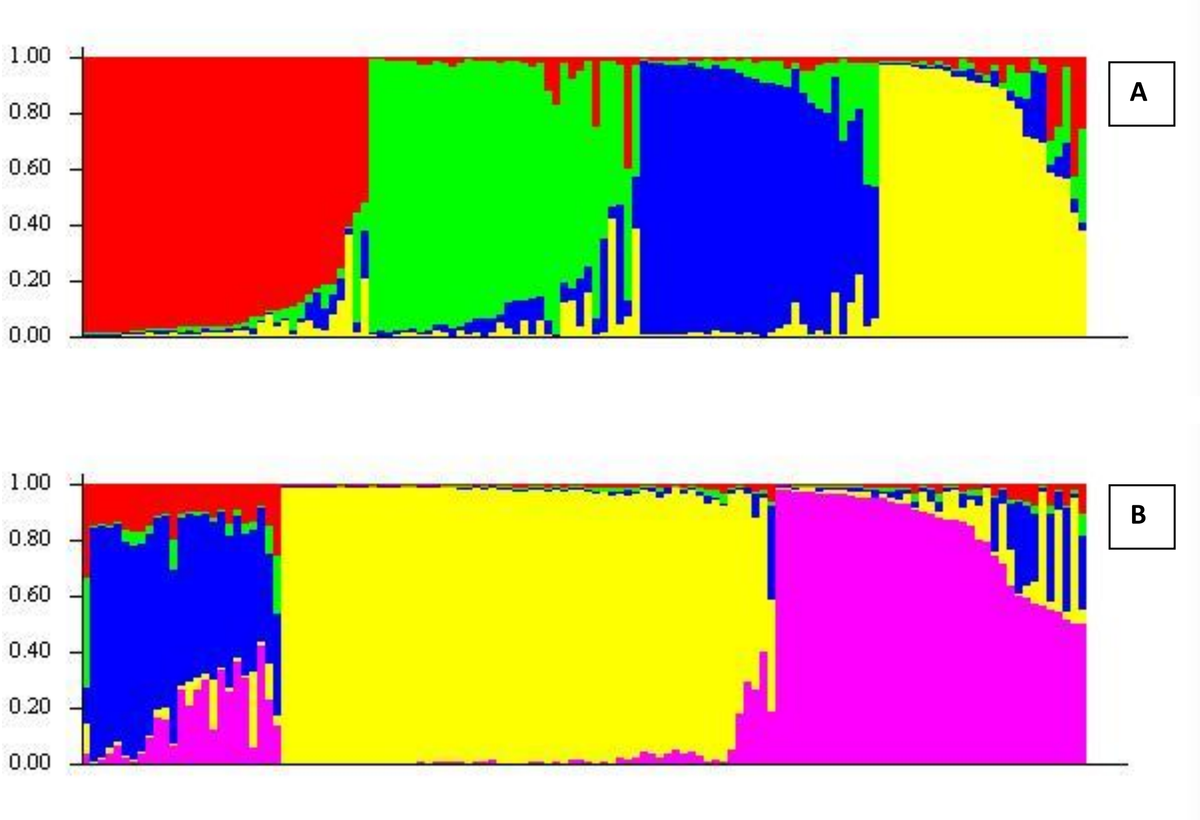


Figure 3.3 Evaluation of the robustness of the grouping by STRUCTURE. Summary structure diagrams are shown for *K*= 4 (Image A), when four clear populations were formed and can be compared with diagram generated when *K*= 5 (Image B), when the populations are not well defined.

The Principal Component Analysis (PCA) method (GENETIX) based on individual microsatellite genotypes revealed no substructure among individual groups and large overlap between individuals from the different populations (Figure 3.4). All four populations exhibited a high degree of allele frequency differentiation. F_{ST} values ranged from -0.297 to 0.704 for individual loci. Mean value for all population was 0.05 (Table 3.3). Only population R2 had an excess of heterozygotes, the other three populations had heterozygote deficiency. F_{ST} values were very low between and within the populations, with highest figure observed in population R1, which means that most of the genetic variance is contained within the populations themselves and also there is a high gene flow between the genetic clusters. The pairwise F_{ST} analysis showed genetic distance between the populations R2 and B, R1 and B but confirmed a close connection of R1 and R2; and R1 and G. This data was further confirmed with a pairwise population matrix of Nei genetic distance and Nei genetic identity (Table 3.4).

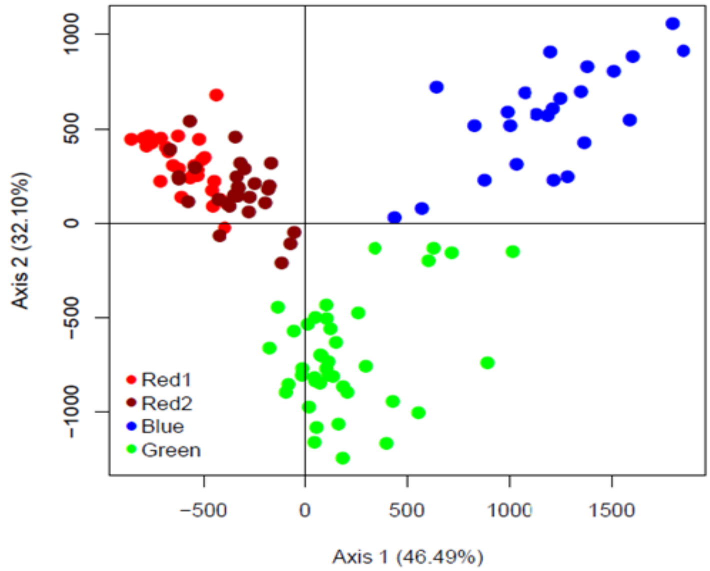


Figure 3.4 Principal Component Analysis (PCA) of four *A. sylvaticus* populations done by GENETIX 4.05 (Belkhir, 2004)

Table 3.3 Results of the genetic differentiation among populations. F_{ST} values were estimated using GENALEX (AMOVA). *Na* = number of alleles, *Np* = number of private alleles. Only the R2 population had a slight excess of heterozygotes, other populations were heterozygote deficient.

Population	*Na*	*Np*	H_e	H_o	F_{ST}
R1	11.4	2.5	0.745	0.652	0.154
R2	10.8	2	0.678	0.684	0.003
B	11.4	2.9	0.81	0.784	0.026
G	13.7	3.7	0.794	0.757	0.049
Mean	11.83	2.3	0.757	0.719	0.058

Table 3.4 Pairwise Population Matrix of *Nei* genetic distance (*Nei* D), identity (*Nei* I) and F_{ST}. The values suggest that all four populations are closely related and probably originated from one source population.

Population		*Nei* D	*Nei* I	F_{ST}
B	G	0.313	0.731	0.044
B	R1	0.510	0.6	0.061
G	R1	0.266	0.766	0.04
B	R2	0.439	0.645	0.072
G	R2	0.250	0.779	0.047
R1	R2	0.145	0.865	0.034

Two different software programs (GENEPOP 3.4, GENALEX 6.502) confirmed that 12 migrants are present in three analyzed populations ($P < 0.01$). Three were from population G, five from R1 and four originated in R2, four of them were adult males, four young female adults, three juvenile females and one juvenile male.

An excess of heterozygotes can suggest that some populations are going through the expansion or early stage of bottleneck. To test this hypothesis the program BOTTLENECK 1.2 (Piry *et al.* 1999) was used. The software compares excess of heterozygosity and heterozygosity expected under drift-mutation equilibrium (Nei *et al.* 1975). In a population at mutation-drift equilibrium (i.e. the effective size of which has remained constant in the past), there is approximately an equal probability that a locus shows a heterozygosity excess or a heterozygosity deficit. A population in mutation-drift equilibrium will have a large proportion of alleles at low frequency (<0.1), and alleles at a low frequency are expected to be more common than alleles present at an intermediate frequency, regardless of the mutation rate (Nei *et al.* 1976, Luikart *et al.* 1998). After a bottleneck there is often a shift in allele frequency distribution towards more alleles being present at an intermediate frequency than at a low frequency. The software calculated the distribution through simulating the coalescent process of *n* genes under two possible mutation models, the Infinite Allele Model (IAM) and the stricter Stepwise Mutation Model (SMM). The allelic diversity is reduced faster than the heterozygosity, i.e. the observed heterozygosity is larger than the heterozygosity expected from the observed allele number were the locus at mutation-drift equilibrium, only if is evolving under IAM (Maruyama and Fuerst 1985). Under the SMM model there can be situations where this heterozygosity excess is not observed (Cornuet and Luikart 1996). However, few loci follow the strict SMM, and as soon as they depart slightly from this mutation model towards the IAM, they will exhibit a heterozygosity excess as a consequence of a genetic bottleneck. The program BOTTLENECK computes for each population sample and for each locus the distribution of the heterozygosity expected from the observed number of alleles (*k*), given the sample size (*n*) under the assumption of mutation-drift equilibrium. Wilcoxon signed rank test showed highly significant heterozygote deficiency in all four tested population under the SMM model (Table 3.5). These results suggest that all four populations experienced a recent demographic expansion instead of bottleneck. Such a deviation can also arise in the earliest stages of a bottleneck, when a transient excess of alleles may exists (Cornuet and Luikart 1996). The mode shift model indicated that the distribution of allele frequencies fitted the normal L-shaped distribution.

Table 3.5 BOTTLENECK 1.2 results (Piry *et al.* 1999) for four tested wood mouse populations in Malham Tarn under the SMM model (Wilcoxon test). H_e = expected number of loci with heterozygote deficiency, H_o = observed number of loci with heterozygote deficiency

Population	***P***	$\boldsymbol{H_e}$	$\boldsymbol{H_o}$
R1	0.005	5.84	1
R2	0.002	5.87	1
G	0.0001	5.9	1
B	0.0015	5.94	1

Approximate Bayesian Computation software (DIYABC v2.1.0) was used (Cornuet *et al.* 2014) to analyse the population history. This program combines large scale simulations and evaluates the summaries of simulated datasets and observed datasets. Seven different evolutionary scenarios were simultaneously tested and the final data compared. The scenario with the highest likelihood was selected and rerun on its own with the same parameters. In this scenario population B was the oldest and original source population, from which derived first the population R2 and later population G. Population R1 was separated independently from the population R2 and has no direct link with population B or G (Figure 3.7). Wood mice reach sexual maturity after two months and they start to produce new offspring after two-four months. Based on this, data can also be used to estimate the time when the divergence of the populations happened. DIYABC calculated that the population R2 separated from the population B approximately 40.6-80.1 years ago. When a new population is created from its ancestral population, there is an initial size reduction for population since the invasive population generally starts with a few immigrants. Based on the allele frequencies, NeEstimator software (Nomura, 2008) estimated contemporary effective population size (Table 3.6). Population B was the largest, followed by R2, while G population had only 55 individuals. These data correspond with the data obtained from BOTTLENECK and DIYABC. Population B is the oldest, ancestral population and later R2 and G were separated. A significant association of mouse genetic group with trapping location was found ($\chi 2$ = 46.6; D.F. = 9; P = 0.000). When we look at the Malham Tarn sampling sites (Figure 3.5), the ancestral B population is located in the north and the mice from this population were not found in Spiggot Hill. From all collected mice, 24 belonged to this population and the highest number (26.67%) was found in Ha Mire (Table 3.7). This sample size is too small to make any final conclusions. With 40 sampled individuals population G was the largest and mice were found in all four sampling sites (Tarn Fen 26.09%, Tarn Woods 55.36% and Ha Mire 10%). Interestingly, Spiggot Hill was inhabited only by two populations, R1 and R2. When we know, that population B is the ancestral one from which separated R2, and from R2 derived R1, we can conclude that no population originated in this sampling site and the mice sampled mice migrated here, possibly from Ha Mire or Tarn Fen.

The COLONY program was used to infer full- and half-sib relationships, assign parentage, infer mating system (polygamous/monogamous) and reproductive skew in both diploid and haplo-diploid species. This program implements a maximum likelihood method to assign sibship and parentage jointly, using individual multilocus genotypes at a number of codominant or dominant marker loci. Twenty two families of wood mice were identified in this sample collection from the study site. These comprised 4 families of full-sibs and 18 families of half-sibs. Nine families were assigned to the population B, 12 to the population G, 11 to R1 and 10 genetically belonged to R2. Most of the identified families comprised only a few mice (n=2), the highest number of detected siblings was 9 with an average of 4.05 mice per family (the table with detected families and *T. gondii* prevalences is presented in Chapter 4). The likelihood probability values for some siblings were smaller, which suggest they were possible second generation siblings or they shared only one of the parents.

Table 3.6 Effective population size estimated with NeEstimator software (Nomura *et al.* 2008) based on linkage disequilibrium random mating model for four different allele frequencies. N= number of sampled mice, N_e = estimated population size

Population	N	N_e estimate for different allele frequency			
		0.05	0.02	0.01	0+
R1	37	134.6	168.0	319.2	319.2
R2	25	136.2	375.0	180.1	180.1
G	40	55.0	156.9	187.8	187.8
B	24	227.2	infinite	infinite	infinite

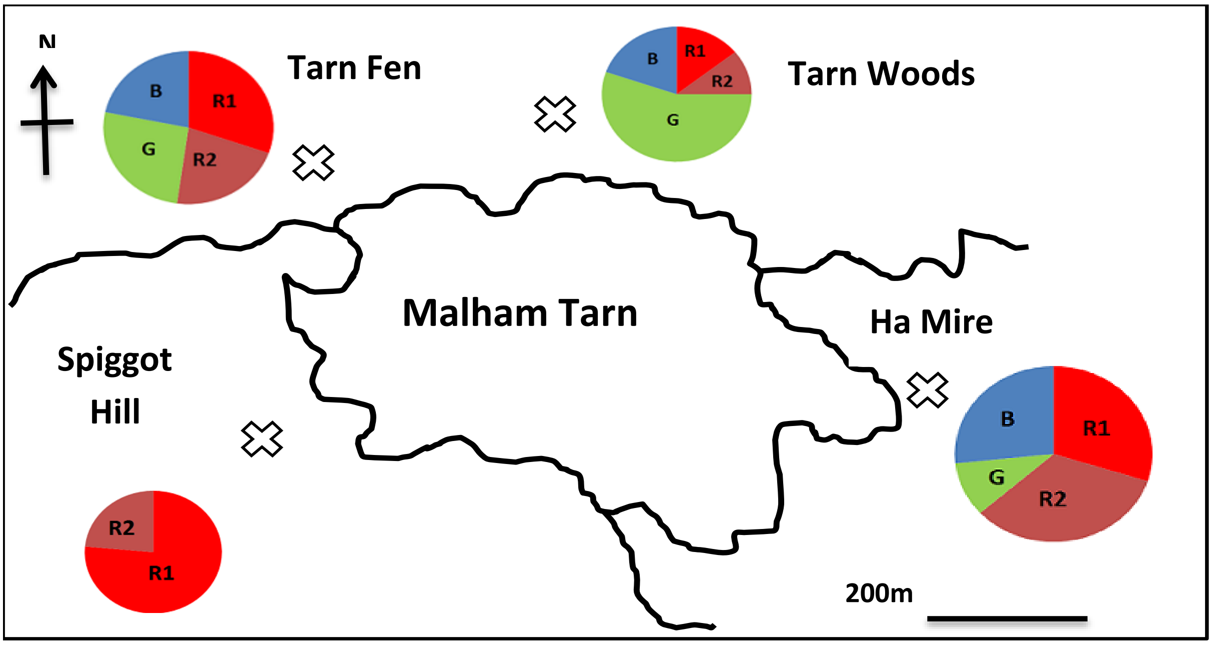


Figure 3.5 The population structure of *Apodemus sylvaticus* in four sampling areas (Spiggot Hill, Tarn Fen, Tarn Woods and Ha Mire) at the study site (Malham Tarn, Yorkshire, UK). The pie charts represent the percentage of each population at each location.

Table 3.7 The population structure of wood mice from the study site (Malham Tarn, Yorkshire, UK) in relation to sampling sites. R, G and B represent the genetic populations as defined by STRUCTURE. Ha Mire, Spiggot Hill, Tarn Fen and Tarn Woods represent the four geographical locations from which the mice were collected (see Figure 3.5).

	R1		R2		G		B		
Site/Population	No.	%	No.	%	No.	%	No.	%	Total
Ha Mire	9	30.00	10	33.33	3	10.00	8	26.67	30
Spiggot Hill	13	76.47	4	23.53	0	0	0	0	17
Tarn Fen	7	30.43	5	21.74	6	26.09	5	21.74	23
Tarn Woods	8	14.29	6	10.71	31	55.36	11	19.64	56
Total	37		25		40		24		126

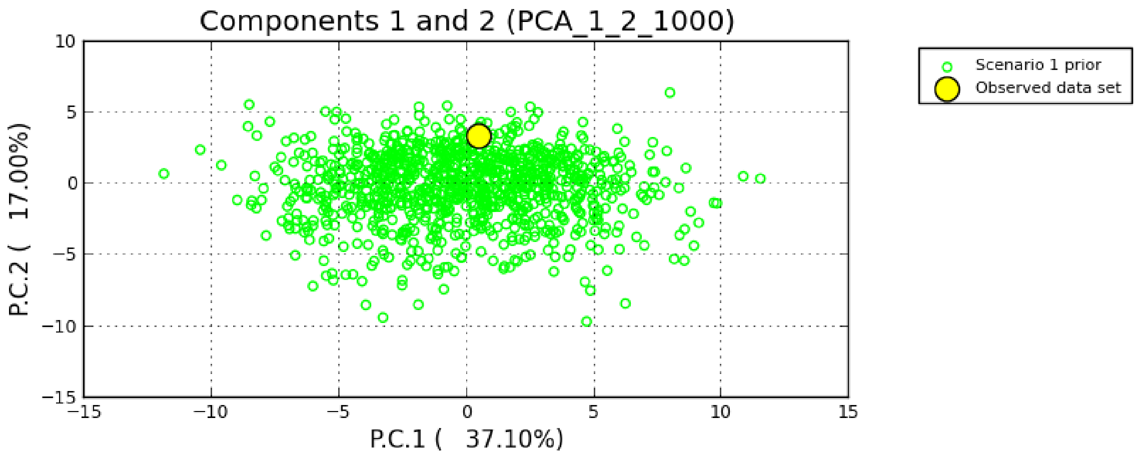


Figure 3.6 Principal Component Analysis (PCA) of four *A. sylvaticus* populations for the final selected scenario in DIYABC (Cornuet, 2014).

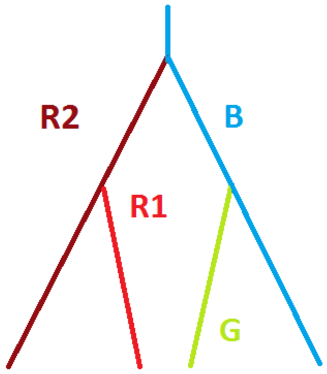


Figure 3.7 Historical scenario supported by highest posterior probabilities in analyses using Approximate Bayesian Computation (DIYABC) and its schematic representation shows population B as an ancestral population from which derived populations R2 and G. Population R1 derived from the population R2.

Data obtained from the GenA1Ex about the population richness and diversity showed very significant (P< 0.001) diversity among the populations (*D*'=0.43, *sH*= 0.233 *O*'=0.594) but very small within individual populations (*D*'=0.818, *sH*=1.922, *O*'=0.182) (Table 3.8). Total Shannon Informational Diversity Index, *sH*(GT) for the wood mouse population from Malham Tarn Nature Reserve was 1.919. This value was within the normal range 1.5-3.5. The value of Simpon's Index of Diversity (*O*'=1-*D*) ranges between 0 and 1 and the greater the value is, the greater the diversity is. This index represents the probability that individuals randomly selected from a sample will belong to different species (category). *D* stands for Simpon's Index and it measures the probability that two individuals randomly selected from a sample will belong to the same species or some category other than species (Simpson, 1949). With this index, 0 represents infinite diversity and 1, no diversity. The greater the value is, the lower the diversity is. The value for the Simpson's Index of Diversity between R1 and R2 was significantly higher than between other populations. It was unexpected to see the highest diversity between these two populations, especially when R1 separated from R2 only recently. For some unknown reasons, the gene flow between these two populations is low and the genetic distance is growing.

Table 3.8 Table of Shannon Pairwise Analysis [0, 1] Scaled Diversity, *D*' and [0, 1], Scaled Overlap, *O*'=1-*D*'

Pop1	Pop2	*D*' Among populations	*O*'=1-*D*' Among populations	Average Within Pops *sH*	Among Pops *sH*
R1	R2	0.276	0.724	1.784	0.145
R1	B	0.552	0.448	1.910	0.301
R2	B	0.500	0.500	1.846	0.285
R1	G	0.412	0.588	1.964	0.230
R2	G	0.387	0.613	1.930	0.211
B	G	0.428	0.572	2.063	0.227
Overall		0.432	0.574	1.922	0.233

3.4. Discussion

This study was designed to provide baseline information on the genetic population structure of the wood mouse (*Apodemus sylvaticus*) using microsatellite analysis. The genetic structure of populations reflects many aspects of their ecology, including past and present demography, interactions with habitat and landscapes, social structure and migration. The results from multiple analytical approaches have shown that the chosen ten microsatellite loci were sufficient for this analysis and the level of genetic variation was very high. The *A. sylvaticus* population in Malham Tarn Nature Reserve showed a high gene flow among different sampling sites and populations. The results indicate that the wood mice from all areas are genetically related but fragmentation of the population is observed and the gene flow might be higher than observed in other studies. All our loci were highly polymorphic (except 8S), with an average number of 23.6 alleles per locus. Mean of the expected heterozygotes was high, averaging from 0.359 to 0.927. Significant departures from Hardy-Weinberg equilibrium were observed on all ten loci. Traditional *F*-statistics, Bayesian clustering analysis and assignment tests indicated the four well defined populations (R1, R2, B and G) are present at four sampling sites. Three localities, Ha Mire, Tarn Fen and Tarn Woods were inhabited by all four populations, only Spiggot Hill consisted of two populations, R1 and R2. The analyses of the population history identified population B as the ancestral population, from which approximately 40-80 years ago (assuming that wood mice start to breed 2-4 months after the birth) separated into R2 and G. There is no evidence, that the divergence of both daughter populations occurred at the same time, but this possibility cannot be excluded. Population R2 subsequently diverged to R1. The B population had the highest number (3.7) of private alleles when compared to other populations. It is surprising that these populations have become genetically differentiated so quickly. F_{ST} values ranged from 0.003 to 0.154 and were significant for all populations except the population R1. Reliability of estimates of genetic summary statistics ($F_{ST,}$ heterozygosity) is more dependent on the number of loci than on the number of sampled individuals. As indicated by a Wilcoxon signed rank test, there was no evidence of an excess of *He* in tested populations. The mode shift test indicated that the distribution of allele frequencies fitted the normal L-shaped distribution expected from a population in mutation-drift equilibrium.

BOTTLENECK data confirmed that all four populations are expanding. Population G had only 55 individuals predicted at the 0.05 allele frequency. This could mean that the expansion is over and the population is now entering a bottleneck, which is characteristic of a sharp reduction in population size. Interestingly, population R2 had a very high population size estimated for allele frequency 0.02 when compared with R1, but not for other frequencies. This can possibly mean that R2 expansion is slowing down, in contrast with R1 where faster expansion can be expected. Infinite values for the ancestral B population suggest that this population is still large enough and probably will continue to grow and expand further.

After a bottleneck, there is often a shift in allele frequency distribution towards more allele present at an intermediate frequency than at a low frequency (Luikart *et al.* 19998). It can be difficult to detect bottleneck in populations of rodents because the methods used in bottleneck detection techniques lack statistical power for the studied population sizes.

In total, 12 migrants were detected, belonging to three populations G, R1 and R2. There was no detected migration from the B population. Migrating mice originated from all four sampling sites and there were no sex differences. The majority of the migrants were adults or young adults. It is possible that a single mating pair may establish a successful population upon entering a new habitat. A small number of effective migrants per generation is sufficient to prevent strong genetic differentiation among populations (Whitlock and McCauley, 1999). Migration can be different between the sexes as there is usually a strong sex bias in dispersal potential in rodents. At the same time, the coexistence of subpopulations connected by a limited amount of gene flow can maintain much more genetic variation across the entire population than a randomly mating population of the same size (Lacy *et al.*1987). The location of the first ancestral B population is a pure speculation but the genetic distribution points at Tarn Woods. The source population must be located close to Ha Mire and Tarn Fen but it must be distant from Spiggot Hill as this site has no individuals from B and G populations. Tarn Woods, a larger forested and compact area is more suitable habitat for different species and can sustain larger populations. The results from NeEstimator showed that population B is the largest one and is not going to reach a bottleneck anytime soon in contrast with population G, for which can be predicted a rapid decrease in population size. In the year 1947 Malham Tarn Field Study Centre was opened. The Centre occupies a large Georgian country house, leased from the National Trust, in 1946 underwent massive reconstruction, two stable blocks in the wood were converted into accommodation and derelict buildings were repaired. The reconstruction of the stables could have forced the wood mice living there before to leave their habitat and migrate to surrounding areas. This time correlates with the divergence time obtained from DIYABC. It is possible that a few mice migrated south (in single or several migrations) and set up population R2, while other mice spread into surrounding woods. Around 1970 experimental blocking of the grip drains on the main bog was begun and this has been pursued systematically since 1987. This change has significantly raised the water table and could have caused the divergence of the R2 population in Spiggot Hill. The Spiggot Hill area is able to sustain a very large population; but is very often flooded during the spring, so probably the mice migrate to surrounding areas. The landscape structure allows easy migration between Ha Mire, Tarn Woods and Tarn Fen area. Gene flow across the landscape additionally suggests that small to medium size roads and grasslands may not be strict barriers to movement of mice. It is possible that the mice migrate (or stay in the original areas) due to high-quality habitats during the low phase, perhaps with source and sink populations and then expand back out into areas of lower quality during the increase and peak phases. A larger forested area and more connected fragments represent a more suitable habitat for different species than smaller, isolated forest fragments. Neighbouring sampling sites Tarn Woods, Ha Mire and Tarn Fen are located within short distance of each other with no obvious barriers and so no isolation by distance is expected in here localized populations. Pairwise Population Matrix of Nei genetic distance (Nei D), identity (Nei I) and F_{ST} values suggest that all four populations are closely related. The fact that gene diversity estimates for the northern (located in Tarn Fen, Tarn Woods and Ha Mire) and southern (Spiggot Hill) populations were relatively high suggests current or relatively recent connectivity. Indeed, pairwise F_{ST} analysis showed that the northern and southern populations are not completely isolated from each other. It is possible that the populations from Spiggot Hill (R1 and R2) effectively act as one large population, but they may have experienced significant fluctuations in size. BOTTLENECK data has shown a faster expansion in R1 population and slower in R2. Although there is clear evidence of ongoing gene flow between neighbouring populations, genetic distance between the populations B (ancestral, source population) and R1 (derived from R2) is growing. The main reason for this is probably the geographical location and spatial dispersion of individual populations. Demographic factors, such as population substructure and social structure can be responsible for significant F_{ST} values. The presence of undetected null alleles can inflate the proportion of homozygotes. It is unlikely that they caused heterozygote deficiency in this study as no null alleles were detected with MICROCHECKER. The significant heterozygote deficiency could be the result of inbreeding or identity by descent. It is possible that mild inbreeding is present in wood mice populations. Genetic drift in small populations can also decrease observed heterozygote frequencies and compound any potential effects of mating among relatives. In contrast, a high degree of polygyny and male-biased dispersal will cause heterozygote excess (Storz *et al.* 2001). The differences in frequencies of heterozygotes between males and females were not significant. Males had a higher number of alleles and lower number of deficient heterozygotes when compared to females. There are only few data on the genetic structure of wood mice populations in natural habitats. Most of the research done on small rodents used different species or the population was located in urban or agricultural areas. This study of wood mice from Malham Tarn Nature Reserve indicates that the level of genetic structure is unusually complex for such a small geographic area, where the distance between individual sampling sites is approximately one mile. For future studies of the wood mice population, a larger sample size is necessary, as a small number of mice can influence the results (i.e. number of siblings in individual families). Ten microsatellites were sufficient enough to determine the population structure, but for the future study more polymorphic marker are recommended. Microsatellite DNA variation reflects the impact of evolutionary forces (mutation, migration, genetic drift and inbreeding) that affect overall genetic variation and the high rates of change at microsatellite loci make them suitable to reconstruct recent changes in population structure (Goldstein and Schlotterer, 1999). On the other hand, microsatellites are neutral and so they don't provide direct measure of the genetic diversity relevant to natural selection and adaption (Frankham *et al.* 2002). As the wood mice have a short life span, more frequent sampling is recommended and it should not exceed the period of two years, as this can affect the population structure. It is important to include ecological and geographical factors into the analysis, as they play an important part as they may drive patterns of genetic diversity. Many of the software programs used in population genetics studies now investigate how ecological factors interact with genetic diversity. Highly variable microsatellite loci are useful genetic markers especially when conducting a pilot study on species, when only a little background information is available.

Chapter 4:

An investigation into the prevalence of *Toxoplasma gondii* infection in wood mice (*Apodemus sylvaticus*) in Malham Tarn Nature Reserve

4.1. Introduction

Toxoplasma gondii is an apicomplexan parasite with a global distribution. The members of the family Felidae are the only known definitive host (Dubey and Frenkel, 1972; Hutchinson, 1965). It infects a whole range of warm-blooded vertebrates, including domestic and game animals, birds and humans (Dubey and Beattie, 1988). There are three main routes of transmission: via oocysts shed in faeces of the definitive hosts, ingestion of tissue cysts and congenital transmission. However, the elucidation of these routes during the past three decades has not elucidated which of these routes is more important epidemiologically. The effect of vertical transmission on the overall spread of the parasite is not clear as its magnitude may depend on the host species, as well as on the parasite strain. The most general mode of spread of toxoplasmosis depends on the presence of the oocysts, which are produced only by the intestinal stages in the cat and the other members of the family Felidae. Experimental studies demonstrate the ability of *T. gondii* to be transmitted congenitally, but few studies have investigated the frequency of this transmission route in natural populations. Murine models are of particular interest when studying the epidemiology of *Toxoplasma* as the cat often predates on small rodents. Experimental infection of both house mouse and the wood mouse with infective oocysts demonstrated a high frequency of vertical transmission (Thierman *et al.* 1957, Beverley *et al.* 1959, Remington *et al.* 1961, Owen and Trees, 1998; Marshall *et al.* 2004, Murphy *et al.* 2008). The results indicated that vertical transmission in *A. sylvaticus* and *M. musculus* is extremely efficient and probably endures for the life of the breeding female. This mechanism favours parasite transmission and dispersion by providing a potential reservoir of infection in hosts predated by the cat. In a natural population of house mice, vertical transmission occurred in 75% of pregnancies , but on the contrary, in BALB/c laboratory mice, and in other species such as rats *Rattus* sp. and hamsters *Mesocricetus auratus,* vertical transmission occurs essentially when infection is acquired during pregnancy, but is less likely in chronically infected individuals (Freyre *et al.* 2006). *Calomys callosus,* a rodent of the family Cricetidae widely distributed in Central Brazil, has been used in previous studies, demonstrating its high susceptibility to *T. gondii* infection (Favoreto *et al.* 1998). Also, this rodent was shown to be a suitable experimental model to study the dynamics of congenital toxoplasmosis, due to the ability of a highly virulent strain of *T. gondii* (RH strain) to infect trophoblast cells during the early blastocyst-endometrial relationship. The seroprevalence of *T. gondii* in rats (*Rattus norvegicus* and *R. flavipectus*) was investigated in

Guangzhou, southern China, between November 2009 and January 2010. In total, 217 rat serum samples were collected; antibodies to *T. gondii* were detected by the modified agglutination test (MAT), and 7 (3.2%) were found positive (Yin *et al.* 2010). Antibodies to *Toxoplasma gondii* were investigated in serum samples of field mice, *Microtus fortis*, from Yuanjiang, Hunan Province, People's Republic of China. The modified agglutination test (MAT) incorporating formalin-fixed whole tachyzoites and mercaptoethanol was used to determine antibodies. Antibodies to *T. gondii* (MAT > or = 1:20) were found in 36 (29%) of 124 trapped mice (Zhang *et al*. 2004).

Toxoplasma was also maintained within the captive rat population in the absence of cats at a prevalence, intensity, and age and sex distribution similar to that of the farmstead rat populations. These results suggest, firstly, that *Toxoplasma* can be perpetuated within wild rat populations without the sympatric presence of cats and secondly, that the congenital route is the predominant route of transmission in wild rats. This study concludes that wild rats represent a significant and persistent wildlife intermediate host reservoir for toxoplasmosis. The prevalence of congenital transmission is low in humans ranging between 0.01 and 1% of live births (Tenter *et al.* 2000), but is much higher in other mammalian species such as sheep (Hide *et* al. 2009; Innes *et al.* 2009). However, this route could explain possible high prevalence of *T. gondii* in some areas with very low cat densities or in marine mammals. In rodents this route of transmission may be more important. Little is known of *T. gondii* genotypes and prevalence in wild populations. The majority of epidemiological studies in wild animals have been based on serological diagnostic methods (Franti *et al.* 1976; Jackson, 1986; Webster, 1994; Dubey *et al.* 1995; Smith and Frenkel, 1995; Hejlicek *et al.* 1997; Jeon and Yong, 2000; Jakubek *et al.* 2001; Yin, 2010). PCR–based studies have shown higher prevalence in some wild mice populations with prevalences ranging from 10.4% (Vujanic *et al.* 2011), 13.6% (Kijlstra *et al.* 2008), 29% (Zhang *et al.* 2004) up to 40.78% (Thomasson *et al.* 2011) and 59% (Marshall *et* al. 2004). The aims of this study were to investigate the detailed distribution of *T. gondii* infection in a series of localised populations of *A. sylvaticus* collected in a systematic manner (Boyce *et al.* 2012; Boyce *et al.* 2013). These populations reside in an area relatively free of cats with density lower than 2.5 cats per km^2 (Hughes *et al.* 2008) but where previous studies have demonstrated a high prevalence of *T. gondii* infection (Thomasson *et al.* 2011). The objectives were to investigate host genotype and spatial location in relation to *T. gondii* infection and genotypes. In this study we investigated the prevalence of *T. gondii* in a wild population of *Apodemus sylvaticus* and combined commonly used *T. gondii* markers and novel tools used in population genetics to investigate the rodent population structure at the family level.

4.2. Materials and Methods

A total of 126 wood mice (*A. sylvaticus*) were captured, using Longworth Traps, from the surrounding area of Malham Tarn Field Centre (Tarn Woods, Tarn Fen, Ha Mire and Spiggot Hill), North Yorkshire between the periods October 2009 - October 2011 (full protocol is described in Chapter 2). The brains were dissected out, using sterile technique, and transferred into sterile tubes containing 400 µl of lysis buffer and stored at −20 °C. Protein was digested and DNA extracted using proteinase K followed by a phenol/chloroform extraction (Duncanson *et al.* 2001). Extracted DNA was tested for mammalian tubulin to ensure viability for PCR (Terry *et al.* 2001). This confirmed the quality of DNA as suitable to undergo PCR, and eliminated the chance of obtaining false negatives due to PCR inhibition.

Detection of *T. gondii* was carried out using nested PCR amplification of the Surface Antigen Genes 1 (Savva *et al.* 1990) as modified by Morley *et al.* (2005). DNA samples were also tested for *T. gondii* by nested PCR amplification of the repetitive and conserved gene B1 (Jones *et al.* 1999). PCR positive brain tissues were used for genotyping using three RFLP markers (SAG2, SAG3 and GRA6) as described by Su *et al.* (2006). All PCR reactions were performed using published primer sequences. All samples were tested a minimum of three times with each marker. Samples that were confirmed to be positive three times were further used for RFLP genotyping with SAG2, SAG3 and GRA6 markers. DNA from sheep known to be infected with *T. gondii* was used as a positive control, sterile water as negative control. All PCR reactions were performed using a Stratagene ROBOCYCLER ™. PCR products were run on 1.5% agarose gel containing GELRED and visualized on an Alpha imager ™ 1220. The risk of contamination was minimized by the use of different locations in each phase of the methodology. For the majority of genotyping reactions, the Type II strain was used as the positive control (as this is the most predominant type in Europe) so that the occurrence of unusual Type I and Type III strains being derived from contamination by the control could be ruled out. To avoid confusion in type calling caused by partial digestion, careful analysis of band sizes, from all markers, was carried out in relation to published marker DNA sequences and other studies.

Microsatellite analysis

DNA extracted from 126 *A. sylvaticus* brain samples was prepared for microsatellite genotyping. All *9 A. sylvaticus* loci have been amplified and genotyped (Makova, 1998). One *Mus domesticus* microsatellite (MS19) was used (full protocol described in Chapter 3). PCR products were then mixed with formamide and LIZ^{TM} standard size marker and genotyped using a GENOTYPER (Applied Biosystems 3130 Genetic Analyser) to gain allele sizes relative to the internal size standard. Samples for each locus were scored with PEAK SCANNER ^{TM} v1.0, that identifies peaks and fragment sizes for application-specific capillary electrophoresis assays. The program *STRUCTURE 2.3.3 (*Pritchard, *et* al. 2000) was used for using multi-locus genotype data to investigate population structure. The program MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004) was used to check for the absence of microsatellite null alleles and scoring errors. Program COLONY v2.0.5.0 (Wang, 2004) was used to estimate the full- and half-sib relationships of mice.

4.3 Results

To examine the prevalence of *T. gondii* in this population of *A. sylvaticus,* a series of PCR reactions were conducted using *T. gondii* specific PCR primers. DNA was successfully isolated from 126 mice brains and tested for the absence of PCR inhibition using amplification of the α-tubulin gene.

Mammalian α- tubulin PCR

All 126 samples showed a high amplification of the α-tubulin gene (Figure 4.1.). This confirmed the quality of DNA to undergo PCR and eliminated the chance of obtaining a false negative due to PCR inhibition. Sheep DNA was used as a positive control for tubulin PCR, sterile water as the negative control.

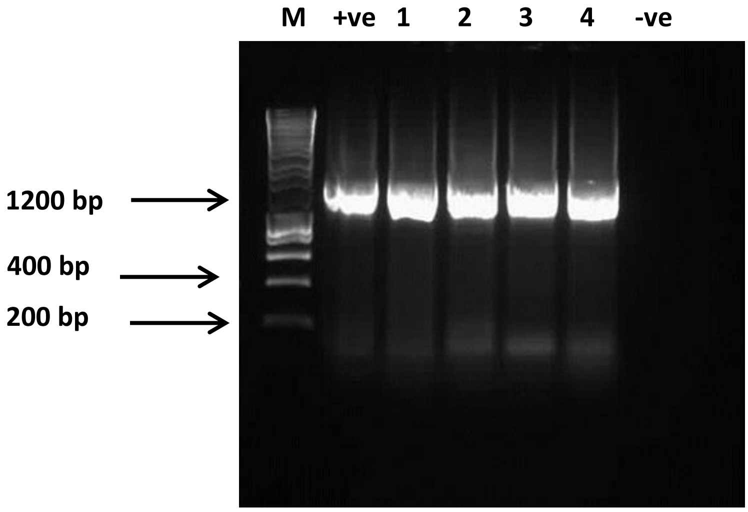


Figure 4.1 Example of mammalian α-tubulin PCR of mice individuals (1-4) on a 1% (w/v) agarose gel. M = Hyperladdder I (Bioline), +ve denotes positive control, *Lanes 3 – 6* = successful PCR amplification in mouse samples, -ve denotes water negative control.

Amplification of B1 Gene

B1 gene was successfully amplified in 12 out of 126 samples (Prevalence of 9.52%, 95% CI: 5.40% - 16.04%). Field samples were considered positive by B1-PCR when second round PCR produced the predicted 96 bp band. The nested reaction did not yield product from either the first or the second-round negative controls. The B1 gene is a 35-fold repetitive gene sequence with unknown function. It is highly conserved which makes it an ideal target for PCR amplification, but this was not confirmed in our case. Other markers were more reliable and showed better amplification of parasite DNA.

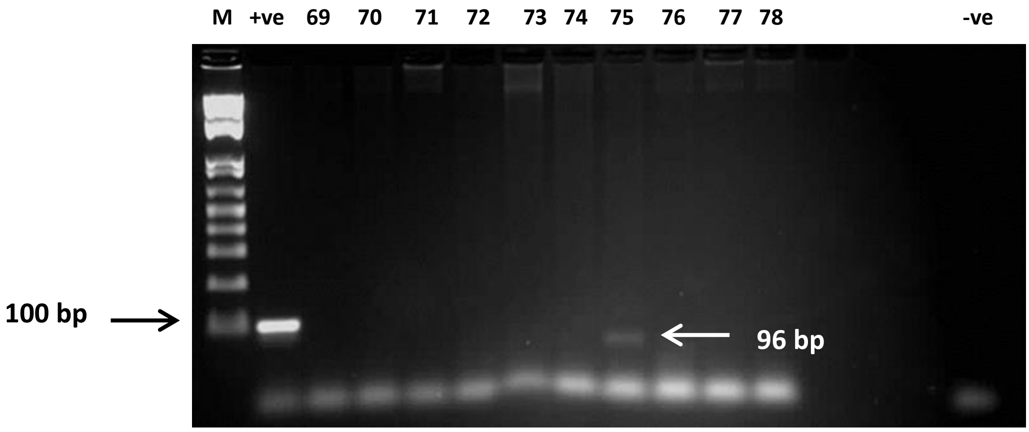


Figure 4.2 An example of B1 PCR 2.5% agarose gel electrophoresis of Malham mice 69-78, M=Hyperladder V Bioline marker, Lane 2 = positive control, Lanes 3-12 = tested samples, Lane 20 = negative control. Sample 75 shows positive amplification of the B1 gene.

SAG1 PCR Amplification

126 brains were available for testing from wood mice. The *T. gondii* SAG1 PCR assay was positive in 44 animals which gave an infection rate of 34.92% (95% CI: 27.1% - 43.6%). An example of successful detection of *T. gondii* can be seen from Figure 4.3. Following the initial first round PCR using the external primers a product of 914 bp is amplified. This product was then used in the second round of the PCR assay to give a highly specific product of 522 bp.

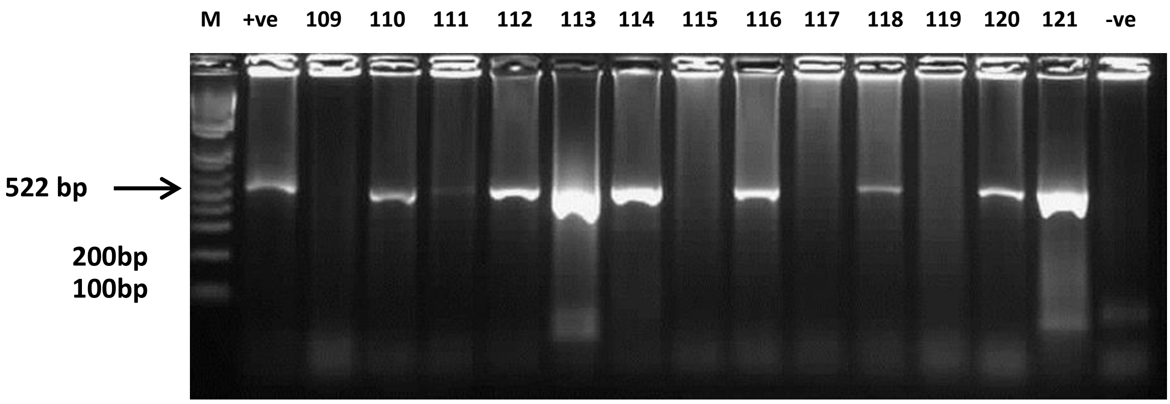


Figure 4.3 SAG1 PCR of wood mice on 1.5% agarose gel. *Lane 1*, Hyperladder 100 bp (Bioline), *Lane 2*, positive control, *Lanes 4, 5, 6, 7,8,10,12, 15* are *T. gondii* positive samples, *Lane 16*, negative control.

SAG2 PCR Amplification

A total of 126 samples were collected and tested for amplification of the SAG2 gene. The SAG2 locus has two polymorphic sites at the $3^{/}$ and $5^{/}$ ends for type II and type III identification (Howe *et al.* 1997). Primers were selected to separately amplify $5^{/}$ and $3^{/}$ ends of the *T. gondii* SAG2 locus as 242 bp and 222 bp products. Any sample found to be positive for only one fragment was re-amplified in order to confirm status, and only samples found positive for both fragments were later genotyped. Amplification of both ends of the SAG2 gene was successful in all 44 positive samples (Prevalence of 34.92%, 95% CI: 27.1% - 43.6%). Negative controls remained free of amplified products.

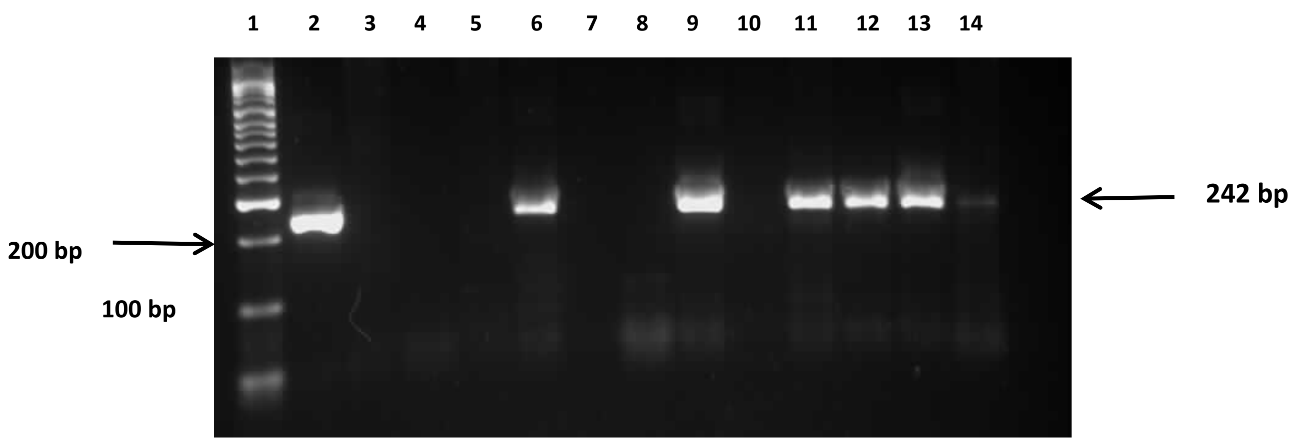


Figure 4.3 Nested-PCR amplification of SAG2 $5^{/}$ end locus, *Lane 1* = Hyperladder II (Bioline), *Lane 2* - positive control, *Lane 3* - negative control. *Lanes* 6, 9, 11, 12, 13 and 14 are positive samples.

Howe *et al.* (1997) previously described a method for the typing of *T. gondii* by restriction analysis of two PCR fragments derived from $5^{/}$ and $3^{/}$ regions of the SAG2 gene, using the enzymes *MboI* and *HhaI*, respectively. In this way, *T. gondii* can be classified as type I (not cut by either *Mbo*I or *HhaI)*, type II (cut by *HhaI* but not by *MboI*) or type III (cut by *MboI* but not by *HhaI*). Nested PCR for $3^{/}$and $5^{/}$ ends of SAG2 gene demonstrated that after digestion of the amplicons with enzymes *MboI* and *HhaI* 16 samples had Type II alleles, 3 had Type III and 18 were mixtures of SAG2 Type II and Type III (Table 4.2). No Type I was present at the SAG2 locus. In four mice we were not able to identify the genotype.

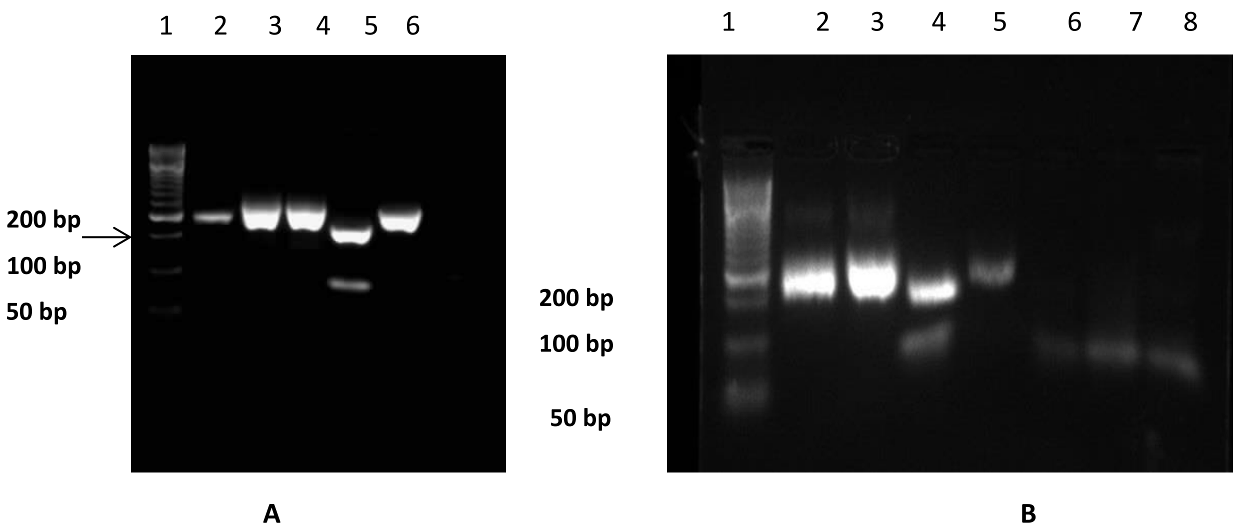


Figure 4.4 Agarose gel electrophoresis analysis of SAG2 PCR amplification products and restriction digests from *T. gondii* infected samples. Image **A** shows PCR-RFLP analysis of the $5^{/}$end flanking region, *Lane 1* is Hyperladder II (Bioline), *Lane 2* is undigested control. Sample in *Lane 5* was digested with *MboI* what indicates Type III. Undigested samples in Lanes 3, 4 and 6 are possible type I or II. Image **B** shows PCR-RFLP analysis of the $3^{/}$end flanking region, *Lane 1* is Hyperladder II (Bioline), *and Lane 2* is undigested control. Samples in *Lanes* 4, 6, 7 and 8 were digested with *HhaI* so that means they are Type II. Samples in Lanes 3 and 5 were uncut and are possible type I or III.

SAG3 PCR Amplification

Genetic analysis of the SAG3 locus was performed to determine the prevalence of the different genotypes of *T. gondii* (strain types I, II, and III) in wood mice from Malham Tarn area. 226 bp products at the end of second round of amplification were considered PCR positive (Figure 4.5). Genetic analysis was performed by PCR-RFLP at the SAG3 locus of *T. gondii*. The amplified SAG3 products were digested with *AlwNI* and *NciI* (full protocol is available in Chapter 2). Unique RFLP patterns generated by these enzymes can be used to identify SAG3 alleles previously associated with Type I, II or III genotypes (examples are presented in Figure 4.6). *NciI* cuts Type I twice (99bp, 65bp and 62bp), Type II stays uncut, Type III cuts once (161 bp, 65bp), *AlwnI* cuts Type II only (128, 98bp); Types I and III stay uncut. Complete characterization of the SAG3 gene was successful in 44 samples (34.92%, 95% CI: 27.1%-43.6%). The digestion of the SAG3 PCR products with *AlwNI* and *NciI* revealed Type III genotype in 4 mice, Type II in 13 mice, 8 mice had the mixture of Type II and Type III, 1 mouse had the combination of Type I and Type III and 16 mice had all three types present at SAG3 locus. Results are summarized in Table 4.2.

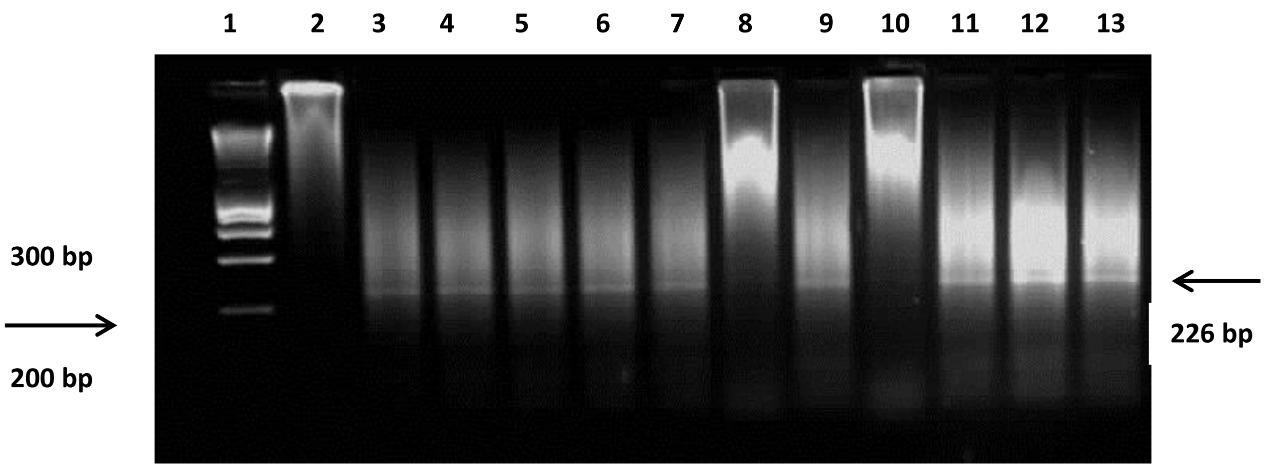


Figure 4.5 PCR amplification of SAG3 locus. *Lane 1*= Hyperladdder II (Bioline), *Lane 2* = negative control, *Lane 3* = positive control, *Lanes 4-13* =positive samples.

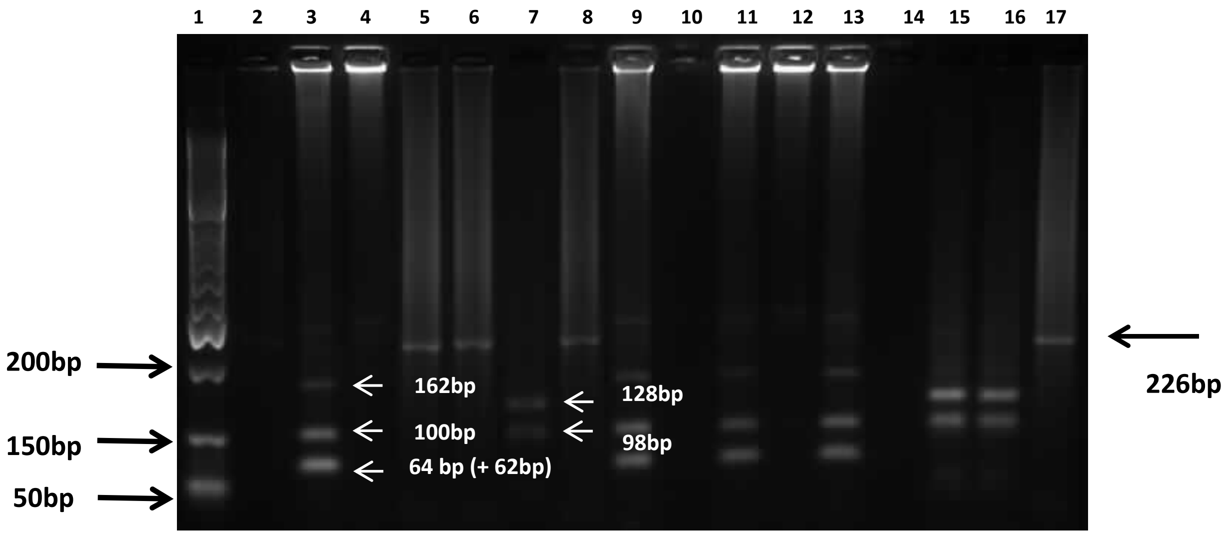


Figure 4.6 Agarose gel (2.5%) electrophoresis analysis of SAG3 PCR amplification products and restriction digests from *T. gondii* infected samples. *Lane 1*- Hyperladder 50bp (Bioline), Lane 3- mixed Type I and Type III positive controls + *NciI* - mixed patterns, *Lane 6* – Undigested SAG3 control (no enzymes added), *Lane 7*- Type II control + *AlwnI*, *Lane 9* – Type I andIII, *Lane 11* – Sample 106 + *NciI* – shows mixed Type I and III, *Lane 12* – Sample 97 + *NciI* – bands disappeared, *Lane 13*- Sample 75 + *NciI* – shows mixed Type I and III, *Lane 15*- Sample 106 + *AlwnI* – Type II present, *Lane 16* – Sample 97 + *AlwnI*- cut Type II present, *Lane 17*- sample 75 + *AlwnI* – uncut- type II not present

GRA6 PCR Amplification

GRA6 nested PCR was performed using published sequences: $F_{1:}$ ATTTGTGTTTCCGAGCAGG, R_1: GCACCTTCGCTTGTGGT and $F_{2:}$ TTCCGAGCAGGTGACC, R_2: GCCGAAGAGTTGACATAG (Khan *et al.* 2005). A 344 bp product at the end of the second round was considered to be positive (Figure 4.7A). For RFLP typing, 5 µl of nested PCR products were treated with *Mse*I in total volume of 20 µl at 37˚C for 1 h and then digested samples were resolved on a 2.5% agarose gel to reveal banding patterns. Complete characterization of the GRA6 gene was successful in 44 samples. Genotyping of the GRA6 locus with *Mse*I revealed that 15 mice had Type II patterns, 14 had Type III and 15 mice contained the mixture of Type II and Type III patterns. No Type I allele was observed at the GRA6 locus (Image 4.7B). A summary of results is presented in Table 4.2.

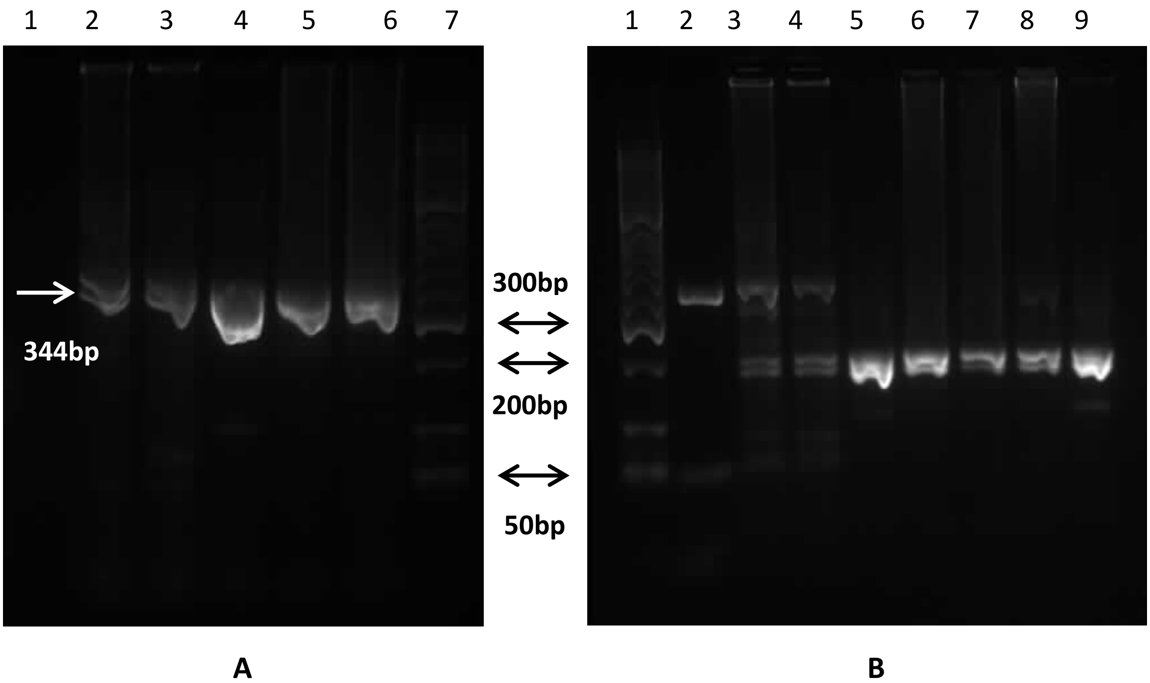


Figure 4.7 Image **A:** PCR amplification of GRA6 locus. *Lane 1* is the negative control, *Lane 2* –positive control, *Lanes 3-6* are *T. gondii* positive samples. *Lane 7* is the 50 bp Hyperladder (Bioline). Image **B** shows PCR-RFLP analysis of GRA6 PCR products. *Lane 1* is the 50 bp Hyperladder (Bioline), *Lane 2* represents undigested control, *Lanes 3, 4* and *9* (very weak bottom bands) show mixed infections of Type II and III (with some partial digests, indicating that possibly Type I is present too), *Lanes 5, 6, 7* and *8* show digest patterns typical for Type II (Khan *et al.* 2005).

Overall prevalence of T. *gondii* prevalence by multilocus PCR

In total forty four samples gave positive reactions with four *T. gondii* specific markers SAG1, SAG2, SAG3 and GRA6. Thus an infection rate of 34.92% (95% CI: 27.1%-43.6%) was found. In total, out of 44 positive mice from Malham Tarn, 5 had Type II alleles at all three tested loci, 23 had the combination of Type II and Type III alleles and 16 had all three genotypes present. Table 4.2 presents the RFLP results for all 44 positive mice using genetic markers GRA6, SAG2, and SAG3 used as described elsewhere (Khan *et al.* 2005; Su *et al.* 2006). Detailed analysis of the RFLP patterns for the three loci showed that for many samples a mixture was the only possible interpretation. A total of 17 juveniles, 65 young adults and 44 adults were present in this cohort of 126 of which 7 (41.18%), 27 (41.54%) and 10 (22.72%) were PCR positive for *T. gondii* (Table 4.1, Figure 4.9).

Table 4.1 The PCR analysis of the brain tissue from *A. sylvaticus* from Malham Tarn. PCR analysis was conducted with 4 *T. gondii* specific markers, SAG1, SAG2, SAG3 and GRA6 (Su *et al.* 2006).

Age Group	Juveniles		Young adults		Adults		Total	
Gender	No. tested	No. positives (%)	No. tested	No. positives (%)	No. tested	No. positives (%)	No. tested	No. positives (%)
Males	7	4 (57.14)	35	12 (34.86)	34	8 (23.53)	76	24 (31.58)
Females	10	3 (30.00)	30	15 (50.00)	10	2 (20.00)	50	20 (40.0)
Total	17	7 (41.18)	65	27 (41.54)	44	10 (22.72)	126	44 (34.92)

Table 4.2 Genotypes of *T. gondii* in wood mice from the study site (NA – marker was not able to be amplified). Genotypes defined as listed previously (Su *et al*. 2006)

	PCR-RFLP Markers		
Mouse No.	GRA6	SAG2 (Combined 5′ and 3′)	SAG3
6	II	II	II
7	III	III	I, II &III
23	III	II &III	II
28	II	II &III	II
31	III	II &III	II
32	III	II &III	II
43	II &III	II	II &III
44	II	NA	NA
52	II	II	II
53	II	II	II
54	II &III	II &III	II
59	III	NA	NA
60	II	II	II
71	III	III	I, II &III
72	III	II &III	II &III
75	II	II &III	I &III
78	II &III	II &III	II &III
84	III	II &III	I, II &III
85	II &III	II &III	II &III
86	II &III	II &III	II&III
87	II &III	II &III	I, II &III
88	II	II	I, II &III
89	II &III	II	II&III
97	II	II &III	II &III
99	II &III	II &III	I, II &III

102	II &III	II &III	III
103	II &III	II &III	III
106	III	II &III	I, II &III
108	II &III	III	II
110	III	III	I, II &III
111	II &III	Not II or III	II &III
112	II &III	III	III
113	II &III	III	I, II &III
114	II	III	I, II &III
115	II	II	I, II &III
116	II	II	I, II &III
117	II	II	II
118	III	II	II
119	II	II	II
120	III	II	I, II &III
121	II	II	III
123	III	II	I, II &III
124	III	II	I, II &III
125	II &III	II	I, II &III

Mice collected in autumn showed higher infection rates when compared with mice from the spring and summer samplings (Figure 4.8). Interestingly, the mice collected in autumn 2011 had almost two time higher prevalences than the mice collected in the same period the year before (Figure 4.8).

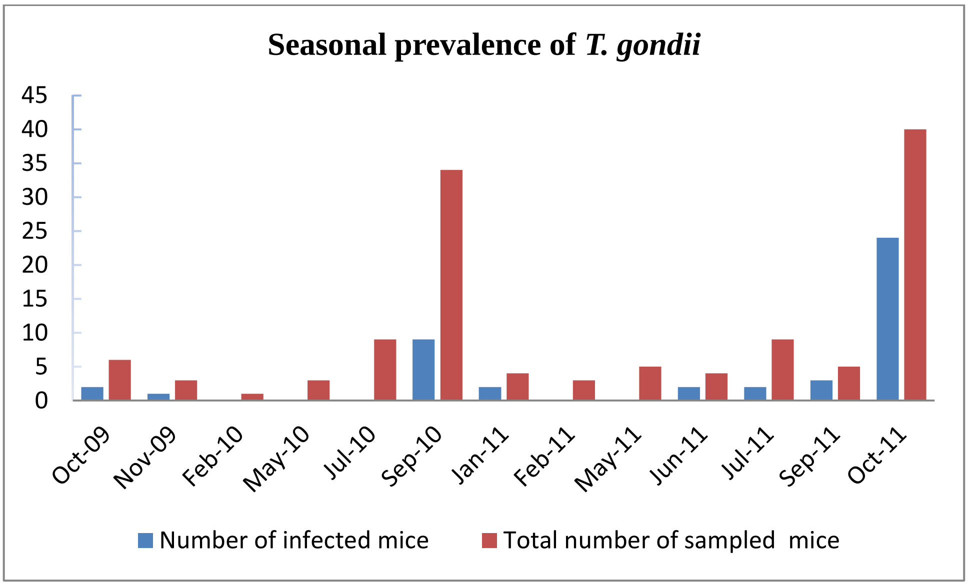


Figure 4.8 Seasonal changes in prevalence of *T. gondii* in wood mice

No significant age prevalence effect was detected in sampled mice, $P = 0.23$ (Figure 4.9).

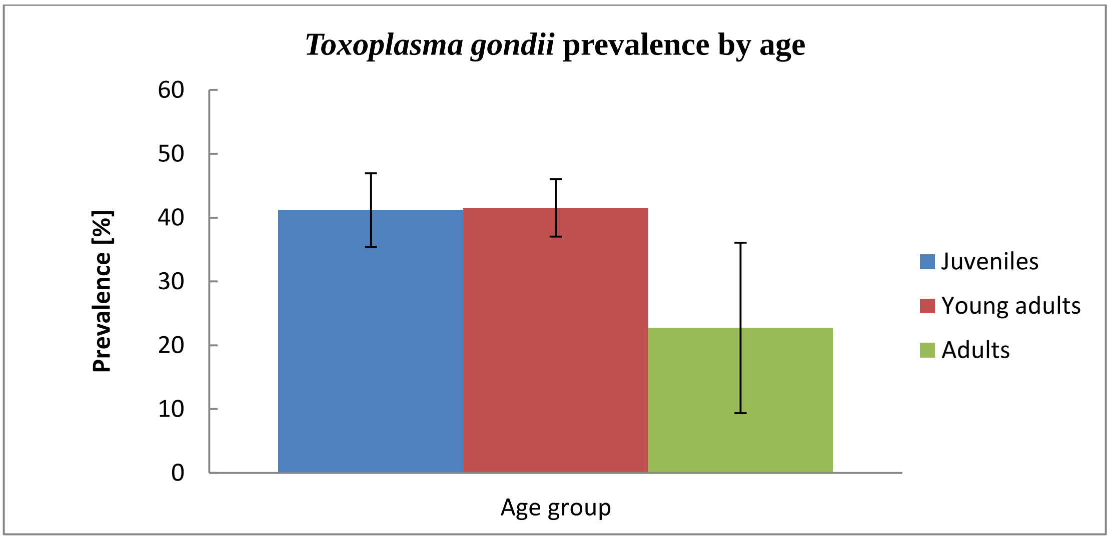


Figure 4.9 The prevalence of *T. gondii* in different age groups of wood mice

A total of 24/76 (31.58%, 95% CI: 22.19%-42.74%) of male and 20/50 (40%, 95% CI: 27.59%-53.84%) of female mice were found to be positive for *T. gondii*. No significant difference was found in prevalence in males and females with $\chi2$ =0.863, D.F. = 1, *P* = 0.353 (Figure 4.10).

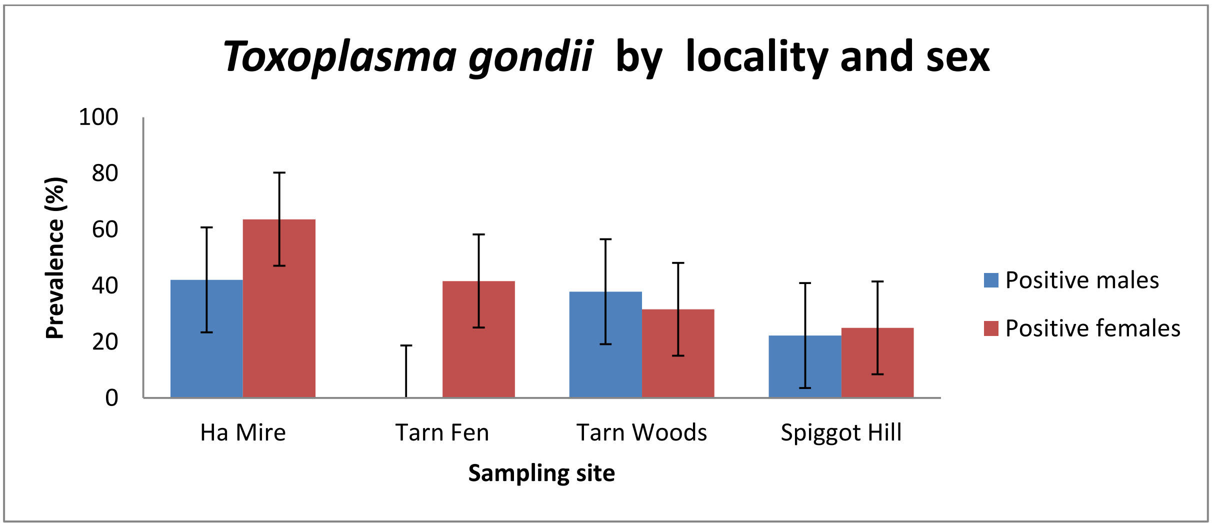


Figure 4.10 The prevalence of *T. gondii* by locality and sex

To understand the distribution of *T. gondii* in mice from different collection sites, the prevalence was calculated for each site (Table 4.3). The null hypothesis of no significant association between prevalence of *T. gondii* infection and location capture was tested. While variation in prevalence was seen at each site, this was not significant (χ^2 = 5.75, D.F = 3, *P*= 0.125). At some sites (e.g. Ha Mire) prevalence that appeared to differ from others, individual pairwise relationships between sites and prevalence were also tested. Only the relationship between Ha Mire and Tarn Fen showed a weak significant association of prevalence with site (Fisher Exact Test, *P* = 0.048) while all other combinations did not.

Table 4.3 Prevalence of *T. gondii* infected *A. sylvaticus* at sites of collection.

Location	No. tested	No. *T. gondii* positives	Prevalence (%)
Ha Mire	30	15	50.00
Tarn Fen	23	5	21.74
Tarn Woods	56	20	35.71
Spiggot Hill	17	4	23.53
Total	126	44	34.92

To understand the distribution of *T. gondii* in mice from different genetic populations, as inferred from STRUCTURE (see Chapter 3), the prevalence was calculated for each population (Table 4.4). Bayesian clustering analysis and assignment tests indicated the four well defined populations (R1, R2, B and G) are present at four sampling sites. Three localities, Ha Mire, Tarn Fen and Tarn Woods were inhabited by all four populations, only Spiggot Hill consisted of two populations, R1 and R2. There was a significant difference across all genetic populations for *T. gondii* prevalence ($\chi 2$ =7.950, D.F. = 2, P = 0.018). The differences in prevalences for some groups were greater than others, so we tested individual for individual pair – wise relationship. Group R1 had a prevalence that showed a highly significant association between each genetic group and prevalence when compared with Group G (Fisher Exact Test (FET), P = 0.003) and Group B (FET, P = 0.0012) but other combinations were not significant (R1 vs R2, FET, P = 0.11; R2 vs G, FET, P = 0.44; R2 vs B, FET, P = 0.15; G vs B, FET, P = 0.6)

Population	No. tested	No. positive for *T.gondii*	Prevalence (%)
R1	37	5	13.51
R2	25	8	32.00
G	40	18	45.00
B	24	13	54.17

Table 4.4 Prevalence of *T. gondii* infected *A. sylvaticus* with respect to genetic population. R1 R2, G and B represent the populations as designated using STRUCTURE. These data suggest that infection status of *A. sylvaticus* is more closely related to genetic population structure than to location of capture.

Some previous studies have suggested that vertical transmission of *T. gondii* might be important in natural populations of mice (Owen and Trees, 1998; Hide *et* al. 2009; Thomasson *et al.* 2011) in which case a relationship should be detectable between infection and family line (Morley *et al.* 2005). To address this question, families were assigned to the mice using the program COLONY and the relationship with *T. gondii* infection examined (Table 4.5). The COLONY program can be used to infer full- and half-sib relationships, assign parentage, infer mating system (polygamous/monogamous) and reproductive skew in both diploid and haplo-diploid species. Twenty two families of wood mice were identified in this collection in the study site. These comprised 4 families of full-sibs and 18 families of half-sibs based, as advised in the program literature, on mice being considered to be siblings if the likelihood probability was more than 0.85. Each identified family was quite small comprising only a few mice (n=2 to n=10). A range of prevalences were found in each of the families, for example, six families had 0% prevalence and one family had 85.7%. The average prevalence in families was 32.55%. This suggests differences in prevalence of *T. gondii* between families and could support the notion of a mechanism that maintains transmission within families. However, this hypothesis was not significant ($\chi 2$ = 0.284, D.F. = 1, P = 0.5943). However this result could be confounded by the very small number of mice in each family. Positive siblings had one, two or all three genotypes present; in some cases the same *T. gondii* genotype was present in all members of the same family.

Table 4.5 Prevalence of *T. gondii* and genotypes present in 22 individual *A. sylvaticus* families assigned by COLONY. Only four families (19-22) are fullsib siblings, where likehood probability was 1.

Family	Number of siblings	Males	Females	*T. gondii* positive	*T. gondii* prevalence (%)	Genotypes Present
1	6	0	6	3	50	II, III
2	4	0	4	2	50	I,II,III
3	7	7	0	4	57.14	II,III
4	3	3	0	2	66.66	II,III
5	5	5	0	2	40	II,III
6	3	3	0	1	33.33	II,III
7	5	0	5	2	40	II
8	3	3	0	1	33.33	II
9	5	1	4	1	20	II
10	7	1	6	6	85.71	I,II,III
11	4	4	0	1	25	II
12	3	0	3	1	33.33	III
13	10	10	0	4	40	I,II,III
14	6	6	0	1	16.67	II,III
15	6	4	2	3	50	I,II,III
16	4	4	0	0	0	0
17	3	3	0	0	0	0
18	3	3	0	0	0	0
19	2	1	1	1	50	II,III
20	2	1	1	0	0	0
21	2	1	1	0	0	0
22	2	1	1	0	0	0

4.4 Discussion

The results of this study demonstrate that *T. gondii* has well-established population in wood mice in Malham Tarn Nature Reserve. We detected a relatively high prevalence of the parasite, with 34.92% (95% CI: 27.14% - 43.59%) of all mice infected in an area with a small number of cats (Hughes *et al.* 2008). This prevalence was lower than the one observed by Thomasson *et al.* (2011) in her earlier study from the same locality, when 40.78% of mice were *T. gondii* positive, but this is not significantly different (P = 0.51). Most reports of *T. gondii* infections in wildlife are from serosurveys. One of the main limitations of serosurveillance is inadequate test sensitivity and so many infections could be undetected. Our results are in contrast with some previously reported data on wild rodents from Germany, when they found 0% prevalence in field mice, shrews and voles (Hermann *et al.* 2012). The prevalences in small mammals in Europe vary, depending upon the species and country where the research was done. In a similar PCR based recent study, 0.7% of *Microtus arvalis* in Austria (Fuehrer, 2010) were positive. High seroprevalences were detected in rodents in Europe in Ardennes (France), the prevalence in shrews (*Sorex* sp.) was 60%, in moles (*Talpa europaea*) 39% and in water voles (*Arvicola terrestris*) 39% (Alfonso *et al.* 2007), in North America in the *Peromyscus* spp., (between 23 – 31 %) while the species *Mus musculus* and *M. californicus* were negative (Dabritz *et al.* 2008). Similar seroprevalence has been detected in wild carnivores, 33% prevalence in polecats (*Mustela putorius*), 18% in stone martens (*Martes foina*) and 17% in pine martens (*Martes martes*), Hejlicek *et al.* (1997), 56% in red foxes (Murphy *et al.* 2007), 43% in artic foxes (*Vulpes lagopus*, Prestrud *et al.* 2008) and 84.7% of foxes (*Vulpes vulpes*) from Saxony-Anhalt and 74.5% of foxes from Brandenburg were seropositive (Hermann *et al.* 2012). These carnivores often prey on rodents, so the prevalence is expected to be higher due to the cumulative effect of consuming infected prey.

Little is known of genotypes of *T. gondii* circulating in wild animal populations. As *T. gondii* genotyping extends to samples from unstudied geographic regions, researchers report an increasing number of atypical strains with newly identified alleles (Dubey *et al.* 2007, Parmeswaran *et al.* 2010). In our study, only five mice had a Type II genotype on all three studied loci, the other mice had two or all three genotypes present. In previous study from this area (Thomasson *et al.* 2011) all three genotypes were detected, only with one difference, that always only one genotype was present in each mouse at the same time. Detailed analysis of the RFLP patterns showed that a mixture was the only possible interpretation. This indicates that mixed infections are more common in natural populations than previously thought. A high rate of mixed infection in intermediate hosts can lead to increased numbers of recombinant genotypes if the final host is present. The mice probably get infected by ingesting oocysts with water or food (insects, grains), by oocyst- contaminated soil or by vertical transmission, either transplacentally or through ingestion of milk.

Numerous studies have demonstrated that *T. gondii* is capable of vertical transmission in a variety of hosts, including humans (Eichenwald, 1948; Thiermannn, 1957; Beverley, 1959; Remington *et al.* 1961; De Roever- Bonnet, 1969; Owen and Trees, 1998; Duncanson *et al.* 2001). Vertical transmission could possibly explain high prevalence in arctic foxes (Prestrud *et. al.*, 2007) where the definitive host is not present and the most important intermediate hosts in this ecosystem, Svalbard reindeers and voles have 0% prevalence and so the presence of oocysts are probably not the main route of transmission. If the infection in Malham Tarn area was caused by oocysts in the environment, we would expect higher prevalence in older individuals, because they had greater opportunity to contract the infection. But in our study there was no significant age prevalence effect and no increased prevalence with age was observed. In contrast adults had lower prevalence when compared with young adults and juveniles. Considering that mice have a relatively short life span, this might seem obvious, unlike in larger animals, where age prevalence can be established. No significant difference was found in prevalence in males and females ($\chi 2$ = 0.863, D.F. = 1, P = 0.353), as higher prevalence would be expected in males if the infection was contracted from the environment, as the males tend to have larger territories and explore the environment more than females. No differences in sex specific prevalence are expected if *T. gondii* infection is transmitted congenitally. Interestingly Tarn Fen was the only site with 0% prevalence in males. Our data suggest that infection status of *A. sylvaticus* is more closely related to genetic population structure. There was a significant difference between each genetic population ($\chi 2$ =7.950, D.F. = 2, P= 0.018). For example, population R1 has the lowest prevalence and has only one individual mouse with a Type I genotype, while the populations B and G have all three genotypes present. Most of the individuals from the R1 population occupy Spiggot Hill and Ha Mire. Spiggot Hill has a much more restricted diversity of genetic groups suggesting that there is less migration to that site. This area lacks trees and is not very suitable to sustain a very large population, indicating that the habitat structure of the landscape allows a slow migration of mice from one population to another, serving as a source population. B and G populations are areas covered with woodlands and tree canopies with high humidity and both have very suitable conditions to sustain large population of rodents. There are several possible interpretations for this. Firstly, different populations have some genetic differences in susceptibility of infection. Population B had the highest prevalence of *T. gondii* (54.17%). This population is the ancestral source population from which about 40-80 ago years diverged R2 and subsequently from it R1 was formed (see chapter 3). Probably, at the same time or later, population G separated from B. Two populations that originated directly from the ancestral one have much higher prevalences than R1. In total 12 migrants were identified in the wood mouse population structures. From these migrants five were *T. gondii* positive, one from population R1 and four originated in R2. It is known that the parasite is capable of altering the natural behaviour of the host to favour its transmission in the environment (Webster, 2001). It is possible that the infected mice from the population B were migrants too which explored more distant habitats and started a new population. As far as we can ascertain, there are no examples of this migration scenario reported, although laboratory studies have shown that inbred lines of rodents differ in their susceptibility to *T. gondii* (Li *et al.* 2012; Gao *et al.* 2015). Previous studies on sheep (Morley *et al.* 2005, 2008) have shown that when pedigree analysis is carried out on family lineages, there is evidence that the parasite is unevenly distributed amongst families even when they are exposed to the same environment. We tried to further analyse these population on family levels but there was no significant difference in prevalence of *T. gondii* between families and the very large 95% confidence intervals suggested that this was almost certainly due to small sample sizes of mice in each family. Most of the families were represented by only two members, the largest was seven. We did see uneven distribution of infection in different families and in some families the same genotypes were present across siblings, but this was not statistically significant. To conduct future work on such a very detailed family based analysis like this, on natural populations of *Apodemus*, will involve sampling on a considerably greater scale.

Nested PCR is more sensitive than conventional PCR but is not preferred for routine diagnosis because it requires two PCR amplification steps. This makes the work more intensive and expensive, with a higher risk of contamination, while it is also more time-consuming. The total amount of DNA from the pathogen will vary among samples in routine practice. One of the major difficulties in the detection of *T. gondi* in animal tissues is the limitation of sample size, as the parasite may be present in the unexamined tissues.

The majority of the previously reported studies analysed only strains from a small fraction (i.e. humans and domestic animals). A sample set that is highly biased towards a certain host species, towards cases of disease or both increases the likelihood of oversampling particular genotypes. It is also possible that limitations in the discriminatory power of the markers used in these early studies missed much of the genetic diversity present. A more diverse population structure was later confirmed as typing techniques with increased resolutions were applied to isolates more representative of *T. gondii's* geographic distribution. A few wildlife samples that were included in Howe and Sibley study (1995) were genetically distinct from the three major clones, with some of them potentially representing recombinants of the three clonotypes. This suggested that a productive sexual stage for this parasite exists in nature and the wider sampling, more representative of *T. gondii's* broad host range, would reveal much greater genetic diversity. However, without sequencing, definitive classification as archetypical strain is not possible in regions where generic diversity is substanti*al*. The dominance of a limited number of genotypes in certain sample sets and from widely geographically dispersed locations is remarkable, but it does not lead to classification as a clonal population (Ajzenberg *et al.* 2002). In some cases the population is highly clonal, in others appears to be epidemic and overall, should be classified as intermediate. To what degree either clonal or sexual propagation can be detected in a sample set is dependent on several factors including the number and genomic location, of genetic markers used, the polymorphism present in the markers, the resolution of the techniques used to detect polymorphism and the host and geographic location of origin of the isolates. It is possible that wildlife samples are perceived to be more genetically diverse than domestic samples because more extensive and sensitive sampling has been applied to the more recently acquired wildlife isolates or a highly clinically biased sample set has been used repeatedly in multiple studies to represent strains in domestic animals and humans.

The role of the cats in transmission is questionable as the infective state occurs for around 2 weeks and therefore all infective oocysts must be released during this period. In this study mixed infections were detected which would mean that a higher number of cats was present in this area or a single cat was infected with multiple strains or a sexual recombination in the cat could generate new combinations of genotypes. There is little evidence to link infection in wood mice to cats, especially when the density of cats is very low. Rodents are unable to acquire the disease by carnivory so the congenital transmission if one of the possible explanation although is generally considered to be relatively rare (Tenter *et al.* 2000). Consequently, *T. gondii* infection levels could theoretically persist within a mouse population in the absence of felid-derived oocysts. The only study to date, to have investigated this route of transmission in a natural mammalian population showed that high levels (61%) of congenital transmission in sheep (Duncanson *et al.* 2001) and urban mice (Marshall *et al.* 2004) occurs. It is clear from the highly clonal structure of *T. gondii* genotypes that the parasite must have succeeded in generating asexual transmission cycles which frequently bypass the definitive host, and this points at transmission in food (Aspinall *et al.* 2001) or vertical transmission (Marshall *et al*. 2004).

Williams *et al*. (2004) detected higher infection rates in Charollais sheep when compared to other pedigrees raising the possibility that pathogenesis varies with host breed or families. Laboratory studies have shown that inbred lines differ in their susceptibility to *T. gondii* (Li *et al.* 2012, Zhao *et al.* 2013). Another explanation is that *T. gondii* is being selectively transmitted through related lineages. Despite the detailed analysis conducted here, it is clear that it will remain difficult to investigate the importance of the different transmission routes of *T. gondii* infection in natural population of wild animals.

Populations R1 and R2 which were found only in Spiggot Hill do not seem to appear separately in any of the geographic regions and it is difficult to estimate how genetically distant is R1 from R2. Spiggot Hill has a much more restricted diversity of genetic groups suggesting that there is less migration to and from that site. Family analysis using COLONY also showed a small number of examples of members of the same family that were trapped in different locations. What is surprising is that there are genetic groups and there must be some mating barriers in this area and they are not based on geographical location as can be seen by the geographical spread of mice from different groups. There might be behavioural barriers to mating (e.g. clans, families) or there might be more complex ecological processes occurring such as annual over winter population bottlenecks followed by spring/summer mice migration. This is a complex ecosystem and there is not enough data to look at such a question. One of the weaknesses is the limitation of the STRUCTURE program when it comes to population number calculations, it is computationally difficult to obtain accurate estimates of Pr(X/K) and the method merely provides an ad hoc approximation and the biological interpretation of *K* may not be straightforward. This is a particular problem for geographically closely related locations where there can be exchange of genetic materi*al.* This will result in an overestimate of population numbers. A further complication for the programme is inbreeding within a population which will also affect population numbers predicted by the programme.

Some scientists have recently presented a hypothesis suggesting that toxoplasmosis is transmitted from infected men to non-infected women during unprotected intercourse (Flegr *et al.* 2014). Arguments for the hypothesis were that tachyzoites are present in seminal fluid and tissue of the testes of various animals, including humans. In some species infection of females by artificial insemination with semen from infected males has been observed. Up to two thirds of *T. gondii* infections in pregnant women cannot be explained by the known risk factors. There is no direct evidence to confirm this hypothesis of sexual transmission but interestingly this hypothesis offers the best explanation for the high number of mixed infections in tested mice.

The prevalence of 34.92 % has been confirmed at Malham Tarn area, with all three genotypes present. The absence of cats in the environment reduces the chances of contracting *T. gondii* via the oocyst route, but still the parasite is perpetuated. One of the possible explanations is congenital transmission. In our study we used a PCR-based diagnostic, which is one of the most specific methods and determines the presence of parasite DNA. Our data showed that the prevalence of *T. gondii* in natural mammal populations is high and the parasite can be perpetuated in the absence of felids possibly via the congenital route.

Chapter 5

Prevalence of infection and genotypic diversity of *Toxoplasma gondii* in patients with lung cancer

5.1 Introduction

Toxoplasma gondii is a protozoan intracellular parasite which can be found in all warm-blooded animals and around 30% of the human population is estimated to be infected (Remington, 1993; Dubey, 2001), with the prevalence reaching 80% in some areas (Tenter *et al.* 2000; Pappas *et al.* 2009). Humans usually get infected from water contaminated with cat faeces, food (ingestion of cysts from undercooked meat), during blood transfusion, organ transplantation or congenitally. In humans, primary infection is usually asymptomatic or does not require medical treatment. If the symptoms occur, they are usually mild influenza like symptoms, rarely hepatosplenomegaly and lymphadenopathy can occur, and are often self-limiting (Krick and Remington, 1978). Latent infections are thought to induce morphological and behavioural changes in infected persons (Webster, 2001). In immunosuppressed and immunodeficient patients, *T. gondii* infection can have fatal consequences. *Toxoplasma gondii* can invade every type of nucleated cell in the body, but target organs such as the lymph nodes, brain, heart and lungs are preferred sites. Proliferation of tachyzoites results in the infection of neighbouring cells and necrosis (Evans *et al.* 1991). Common presentations include encephalitis, pneumonia and myocarditis.

Pulmonary pneumonia is very rare in immunocompetent patients, but has been recognized with increased frequency in patients with AIDS (Caterral *et al.* 1986) and its general frequency is estimated to be between 0.2 and 3.7% but is probably underestimated (Knani *et al.* 1990). In cases from Brazil the mortality amongst AIDS patients due to the pulmonary toxoplasmosis was estimated at 55% (Pomeroy *et al.* 1992) while, for example, the mortality rates of *T. gondii* induced encephalitis in AIDS patients were between 10-40% (Vidal *et*. al 2004). The main problem was that the recognition of *T. gondii* infection in these patients is often delayed. The problem is that the clinical findings often mimic those of other atypical pneumonias. In patients with disseminated toxoplasmosis, 33% had evidence of subclinical pulmonary involvement, even if pneumonia had not been diagnosed clinically. The patients with pulmonary toxoplasmosis could be divided into two groups, patients which present clinically apparent pneumonia as the main clinical feature and patients, where no evidence of pneumonia was observed but other organs were involved (Pomeroy *et al.* 1992). The clinical and radiological features are nonspecific (Cohen *et al.* 1985). Serological studies lack sensitivity and specificity and are of limited value (Montoya, 2002). Definitive diagnosis of toxoplasmosis requires either histological demonstration of the organism or isolation by culture. *Toxoplasma gondii* can be detected by haematoxylin and eosin staining, but, for example, in one study of cerebral toxoplasmosis, the organism was not seen in 50% of positive cases (Luft *et al.* 1984).

Patients with cancer may have deficient cellular immunity and are potentially more susceptible to *T. gondii* infections. Not much attention has been paid to this group of patients and only a few reports are available. These are mostly based on serological studies and not much is known about the genotypes that are present. In a study done by Yuan and colleagues (2007), the seropositivity rates of *T. gondii* IgG in nasopharyngeal carcinoma (46.15%) and rectal cancer (63.64%) groups were significantly higher than the other cancer groups, pulmonary carcinoma (4.55%), breast cancer (9.53%), gastric carcinoma (10.00%), hepatocellular carcinoma (14.29%) and uterine cervix carcinoma (12.50%). They proposed that this demonstrated that there is a likely association between *T. gondii* infection and some kinds of cancer. Interestingly, the IgM rates were not statistically significantly higher. Immunoglobulin M (IgM) is the first immunoglobulin to appear in the immune response. Immunoglobulin G (IgG) is the predominant antibody in the secondary immune response which is most commonly seen with recurrent exposure to specific antigen. If IgG and IgM are both negative, this indicates the absence of infection or extremely acute recent infection. If testing reveals a positive IgG and negative IgM, this indicates an old infection (older than one year). If both IgG and IgM are positive, this indicates either a recent infection or false positive test results (Liesenfeld *et al.* 1997). The rates in Liesenfelds' study were lower than those described in other reports (Yazar *et al.* 2004). Thomas and Lafferty (2011) compared national figures from 37 countries and found that brain cancers are more common in countries where *T.gondii* infections are more common. Brain cancer was 1.8 times more common in countries where *T.gondii* was most common, than in those where it was virtually absent. To take the two extremes, in South Korea, where 4 % of people are infected, 3.4 people out of every 100 000 develop brain cancer each year. In Brazil, 67 % of people are infected, and 5.5 out of every 100 000 develop brain cancer each year. Previous investigations have revealed that *T. gondii* could cause gliomas in experimental animals (Wrensch *et al.* 1993). Studies carried out by Ryan *et al.* (1993) showed that antibody positivity to *Toxoplasma* is associated with meningioma. However, molecular events involved in *Toxoplasma* induced brain cancers are not well understood. The team of Dr.

Gnanasekar (Thirugnanam *et al.* 2013) are testing the hypothesis that micro RNAs (miRNAs) are specifically associated with brain cancers. They think that *T. gondii* is capable of manipulating host miRNAs, which play a central role in post-transcriptional regulation of gene expression (work still in progress).

Traditional studies have concluded that majority of *T. gondii* strains comprises three distinct clonal lineages designed I, II and III (Howe *et al.* 1995; Grigg *et al.* 2001). The parasite appears to reproduce largely clonally in nature, with sexual recombination occurring only rarely. Several reports show that type II is the most prevalent genotype in Europe and North America (Darde *et al.* 1997, Sibley *et al.* 1992). Type II has been also frequently associated with human diseases (Howe and Sibley, 1995), when more than 70% of human cases of toxoplasmosis were found to be caused by this genotype. It is well known that the virulence of the parasite differs in laboratory animals, depending on the *T. gondii* strain (Johnson *et al.* 1997). Type I strains are highly virulent in mice and generate higher levels of parasitaemia than types II and III (Howe and Sibley, 1995). The type I strain in humans has been linked to a high frequency of ocular toxoplasmosis and congenital toxoplasmosis, while SAG2 Type I predominated (75%) among other cases of congenital infection (Fuentes *et al.* 2001). Type II has been often associated with toxoplasmosis in AIDS patients with some fatal consequences (Fuentes *et al.* 2001). Type I and atypical strains have never been reported in asymptomatic or benign congenital toxoplasmosis (Ajzenberg *et al.* 2002). Type III is not very common in the human population, but Grigg (2001) observed atypical recombinant I and III genotypes in patients with ocular toxoplasmosis and AIDS. Aspinall (2002) in her study on patients from England and Wales discovered type III in only one case out of 32 and neither of the genotypes seemed to be associated with particular disease. She also found no correlation between the SAG2 type and clinical presentation, but detected an unusually high number of mixed SAG2 I and II genotypes in the PCR products. In study performed on amniotic fluid, only one case of mixed infection (Type I and II) was detected by direct typing but interestingly murine inoculation of the sample led to only one Type II strain isolation (Villena *et al.* 2004). It is possible that atypical/recombinant genotypes are associated with severe toxoplasmosis at least in immunocompetent patients (Carme *et al.* 2002). Reactivation of unusual I/III genotype in HIV positive patients in France (originating from Ghana) led to severe encephalitis and chorioretinitis, followed later by septic shock, organ failure and death (Genot *et al.* 2007). All these studies were based mostly on SAG2 PCR-RFLP (Howe *et al.* 1997) and used old *T. gondii* nomenclature using only a few markers. The main problem with this system is that a single allele does not necessarily imply that all alleles are of the same type and it makes it very difficult to identify atypical strains. Multilocus genotyping is necessary to reveal atypical or recombinant genotypes. New multilocus nested PCR designed by Su *et al.* (2006) and reported in Swab *et al.* (2013) uses 10 *T. gondii* markers able to detect a much greater variety of genotypes (SAG1, SAG2, SAG3, BTUB, GRA6, L358, APICO, PK1, c22-8 and C29-2) and all known RFLP genotypes are collected in the *T. gondii* database (www.toxodb.org). Analysis with these new markers can provide more data on *T. gondii* population structure and reveal possible link between some unusual genotypes and diseases in humans and animals.

The aims of this study were to determine the prevalence and genotypes of *T. gondii* in humans with lung cancer and a variety of lung diseases (asthma and COPD). Parasite prevalence will be determined by parasite specific PCR, histological and immunochemical methods. Due to the nature of the available human tissue samples, identification of parasite genotypes will be conducted using PCR and genotyping directly from the samples.

Objectives

- To extract DNA from lung tissue obtained from cancer patients
- To confirm the quality of DNA extracted using mammalian tubulin PCR
- To detect the presence of *T. gondii* using nested PCR amplification of *T. gondii* specific genes (B1, SAG1, SAG2, SAG3)
- To detect the presence of *T. gondii* using haematoxylin and eosin staining
- To detect the presence of *T. gondii* with *T. gondii* specific Immunohistochemical staining (using commercial antibodies)
- Identify *T. gondii* genotypes in positive samples using RFLP on *T. gondii* specific genes (SAG1, SAG2, SAG3, GRA6, Apico, L358, alt.SAG2, BTUB, c22-8, c29-2, PK1 (Su *et al.* 2006, Shwab *et al.* 2013)
- To analyse infection data in relation to biological range of known clinical parameters (age, sex, disease etc.)

5.2 Materials and Methods

In total, 88 DNA and 76 tissue samples from lung cancer patients from Wythenshawe Hospital were tested for the presence of *T. gondii* using molecular, histological and immunohistochemical methods as described in Chapter 2. The tissue was collected form non-cancerous tissue taken as a part of lung resection for removal of cancerous tissue. From all 88 patient samples, 12 had some missing data or tissue was not available for immunohistochemistry so these samples were excluded from the final analysis. The presence of the parasite was tested with nested PCR with five different markers, SAG1, SAG2 ($3^{/}$ and $5^{/}$ were tested separately), SAG3 and B1 (as described in Chapter 2). Each DNA sample was tested three times at its original concentration. Samples were accepted as positive if the presence of the parasite DNA was confirmed in all three reactions. Ten additional *T. gondii* genes SAG1 (Su revised method), SAG2, SAG3, GRA6, Apico, L358, alt.SAG2, BTUB, c22-8, c29-2, PK1 (Su *et al.* 2006, Swab *et al.* 2013) were used as RFLP markers for direct genotyping of parasite DNA from PCR positive lung tissues. With these markers, samples were also tested three times and were accepted as positive if at least one reaction gave positive results. If samples were negative in all three reactions they were considered to be negative with that marker. Two negative controls were used, sterile water and DNA from a wood mouse negative for *T. gondii*. A Type II strain from Slovakia and RH strain (Type I) were used as positive controls. Each marker was tested on *T. gondii* negative human genomic DNA (Promega, Madison, WI, USA, product code G304A) to check for false positive results. Immunohistochemistry was performed in paraffin embedded tissue using commercial polyclonal antibodies produced in rabbits (Thermo Fisher Scientific, Catalogue number PA1-38789, Rockford, IL, USA). The tissue sections were cut in 5 µm sections and mounted on positively charged glass slides. Three negative controls were used, lung sections from *T. gondii* negative wood mouse (*Apodemus sylvaticus*), human lung sections with primary antibodies omitted and cells derived from a C2C12 culture (mouse myoblast cell line, free of *T. gondii*). Cell culture derived *T. gondii* RH strain tachyzoites and lung tissue from a *T. gondii* infected wood mouse were used as positive controls. Immunostained slides were assessed using quantitative criteria (ImageJ) to provide a percentage score which described the degree of infection (as described in Chapter 2). Finally, Haematoxylin and Eosin staining was used to confirm that structures compatible with *T. gondii* stages can be observed within sections. Data was analysed by logistic regression. All analyses were undertaken using R (R Core Team, 2013).

5.3 Results

5.3.1 Molecular analysis of the human lung tissue samples

The tissue samples were collected from 88 patients with lung cancer at the Wythenshawe Hospital, Northwest Lung Centre in Manchester as a part of long term study on lung disease. For each sample, data was available on age, gender, lung conditions (COPD, asthma) if present and on smoking status of individuals. For 12 patients, not all data or tissue of suitable quality was available, so they were excluded from the final analysis but we included them in genotype studies and using the molecular diagnostic markers. DNA was successfully isolated from 88 frozen lung tissues blocks and tested for the absence of PCR inhibition using PCR amplification of the α-tubulin gene. All 88 samples showed good amplification of the α-tubulin gene and were used for further testing (Figure 5.1). PCR products from B1, SAG1, SAG2 3', SAG2 5' and SAG3 were sequenced to confirm the correct identity of PCR amplified bands (Figure 5.20, 5.21, Appendix).

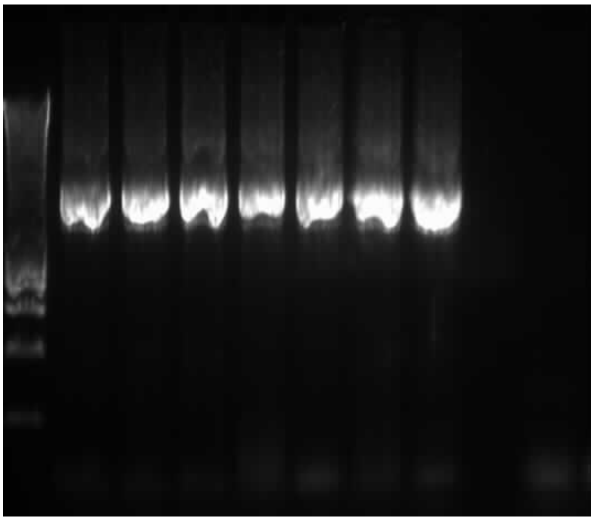


Figure 5.1 PCR amplification of the α-mammalian tubulin gene to check for PCR inhibition *Lane 1* = Hyperladder 100kb (Bioline), *Lanes 2-7* = tested samples, *Lane 8* = positive control (woodmouse DNA), *Lane 9*= negative control. Expected product size was 1600bp.

We used highly sensitive nested PCR protocols to detect *T. gondii* in the group of lung cancer patients. DNA extracted from 88 frozen lung tissue blocks was tested for the presence of the parasite with four markers B1, SAG1, SA2 ($3^{/}$ and $5^{/}$ end were tested separately) and SAG3.

Each sample was tested three times at the original DNA concentration. The samples were considered to be positive if they successfully amplified in all three reactions. Negative controls (water and *T. gondii* negative woodmouse DNA) were interspersed throughout the PCR reactions to detect any false amplification. Different stages of the PCR assay were conducted in separate rooms to minimize the risk of contamination. PCR products were visualized by agarose gel electrophoresis. All 88 tested samples (36 females and 52 males) were confirmed to be positive for *T. gondii* with markers B1, SAG1, SAG2 3$^{/}$, SAG2 5$^{/}$ and SAG3, giving a prevalence of 100% (95% CI: 95.19% - 100%). Examples of the gel images for tested markers can be seen in Figures 5.2-5.5.

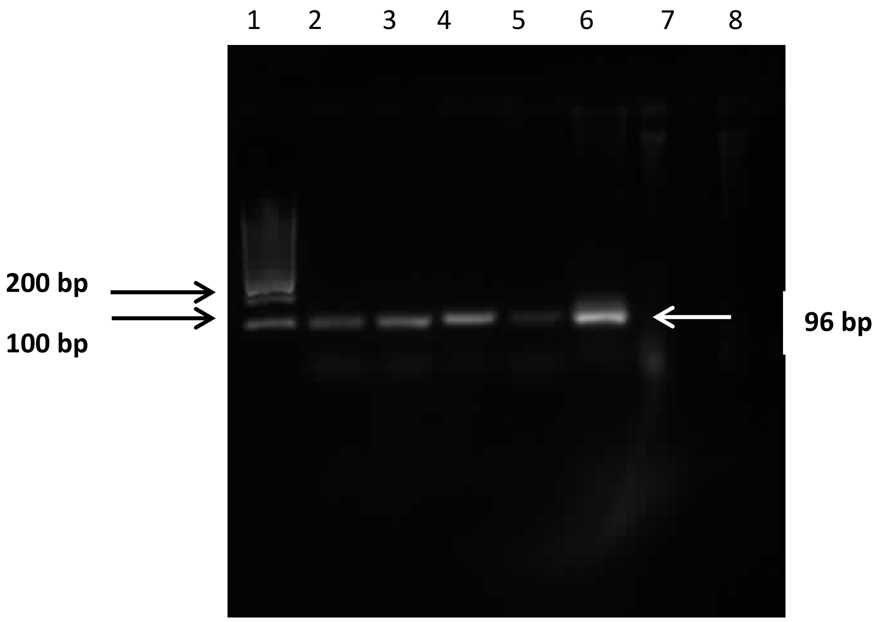


Figure 5.2 An example of B1 PCR gel electrophoresis on human lung cancer patients. *Lane 1* is the Hyperladder 100bp (Bioline), *Lanes 2-5* are *T. gondii* positive patients, *Lane 6* is the positive control (expected 96bp product), *Lanes 7* and *8* are the negative controls (sterile water and *T. gondii* negative wood mouse).

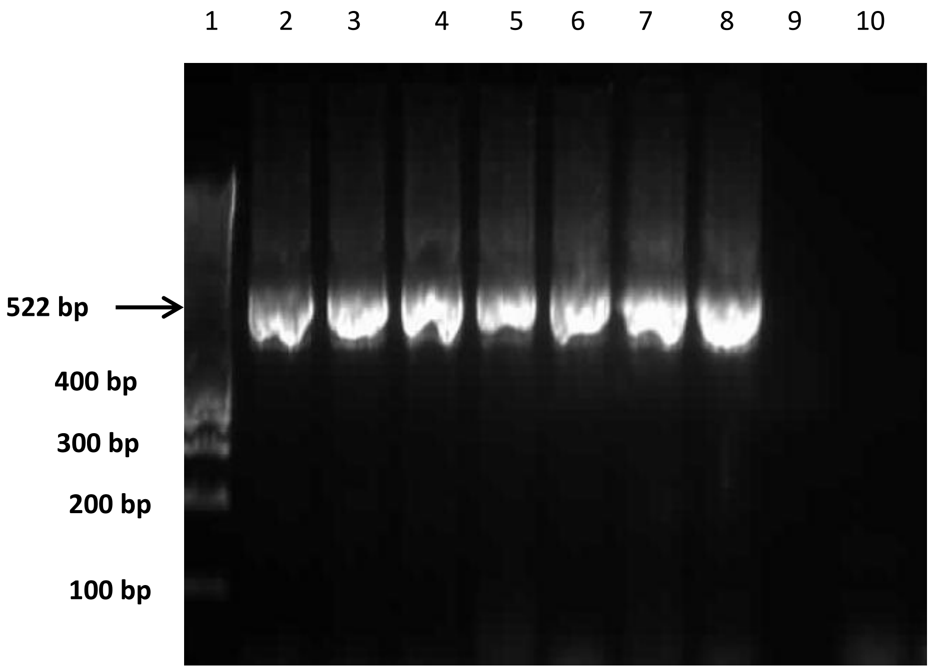


Figure 5.3 Example of SAG1 PCR of lung tissue from human lung cancer patients on 1.5% agarose gel. *Lane 1,* Hyperladder 100 bp (Bioline), *Lane 2* = positive control (522bp product size), *Lanes 3, 4,5,6,7,* and *8* are *T. gondii* positive patients, *Lanes 9* and 10 are negative controls.

SAG2 locus has two polymorphic sites at the $3^{\prime}$ and $5^{\prime}$ ends for type II and type III (Howe *et al.* 1997), so amplification of this locus was performed separately. Two nested PCRs, one targeting the $3^{\prime}$ end of the SAG2 gene and the other targeting the $5^{\prime}$ end of SAG2 gene were optimized as per the method of Fuentes *et al.* (2001).

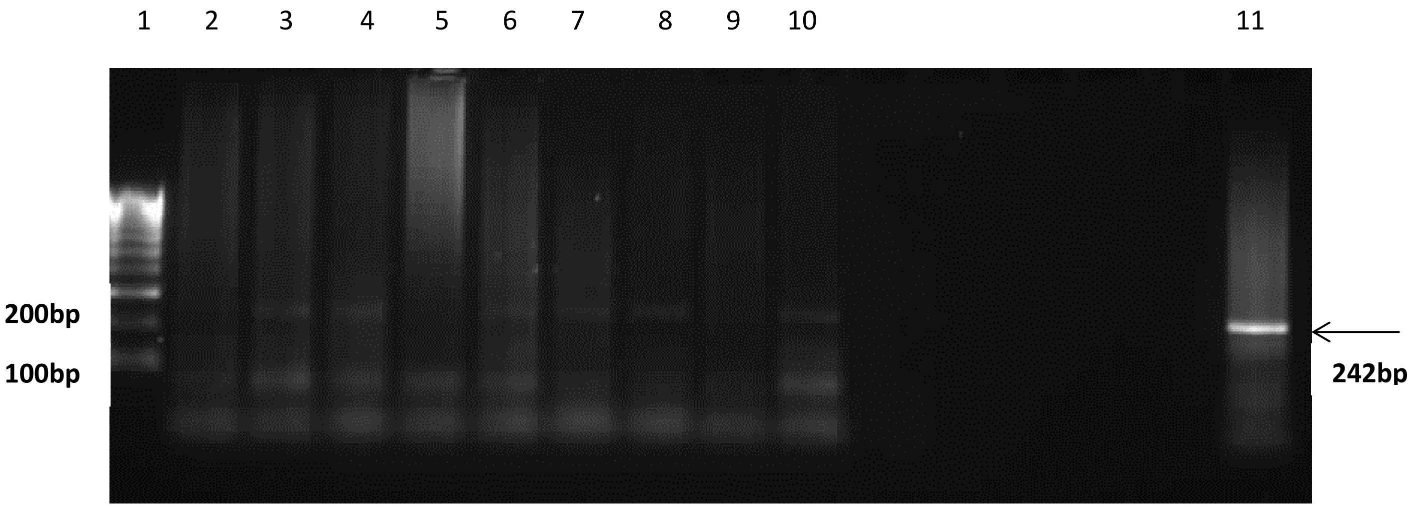


Figure 5.4 An example of nested-PCR amplification of SAG2 $3^{/}$ *Lane 1* = Hyperladder 100bp (Bioline), *Lane* 5 is the negative control, *Lanes 1,2,3,4,6,7,8,9* and *10* are positive patients (some very weak bands are present), *Lane 11* is the positive control (expected product size 242 bp). The results were clearer on original gels by adjusting exposures and on computer version but on printed images it is difficult to see the bands.

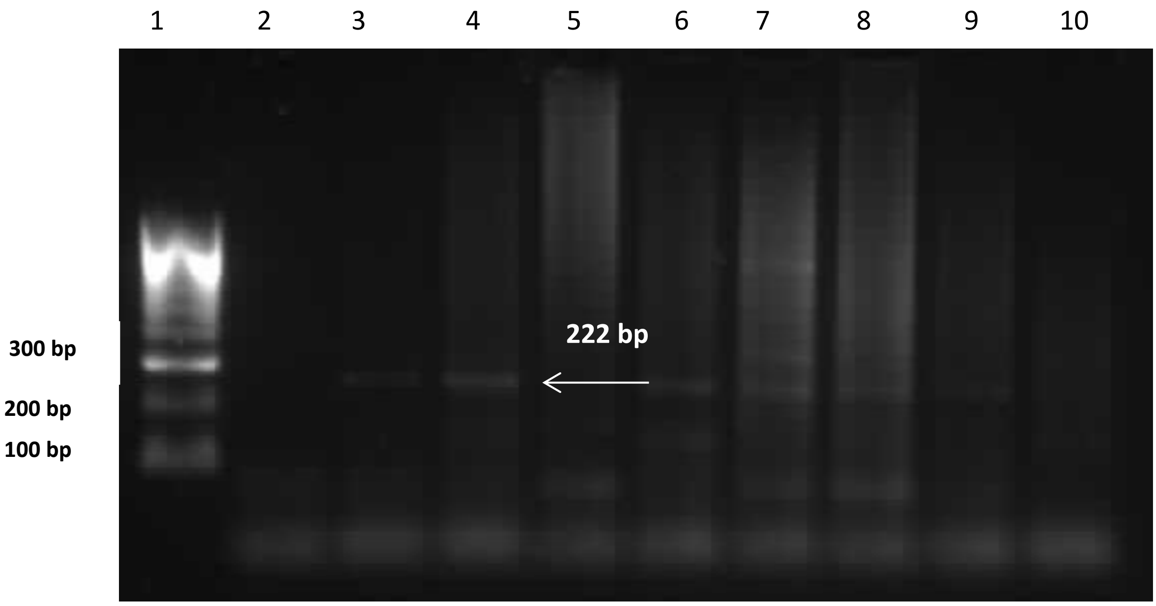


Figure 5.5 An example of nested-PCR amplification of *T. gondii* SAG2 $5^{/}$ end locus from lung cancer patients. *Lane 1* = Hyperladder 100bp, *Lanes 2* and *10* = negative controls, Lane3 = positive control, Lane 4, 5, 6, 7, 8 and 9 = positive patients. The results were clearer

on original gels by adjusting exposures and on computer version but on printed images it is difficult to see the bands.

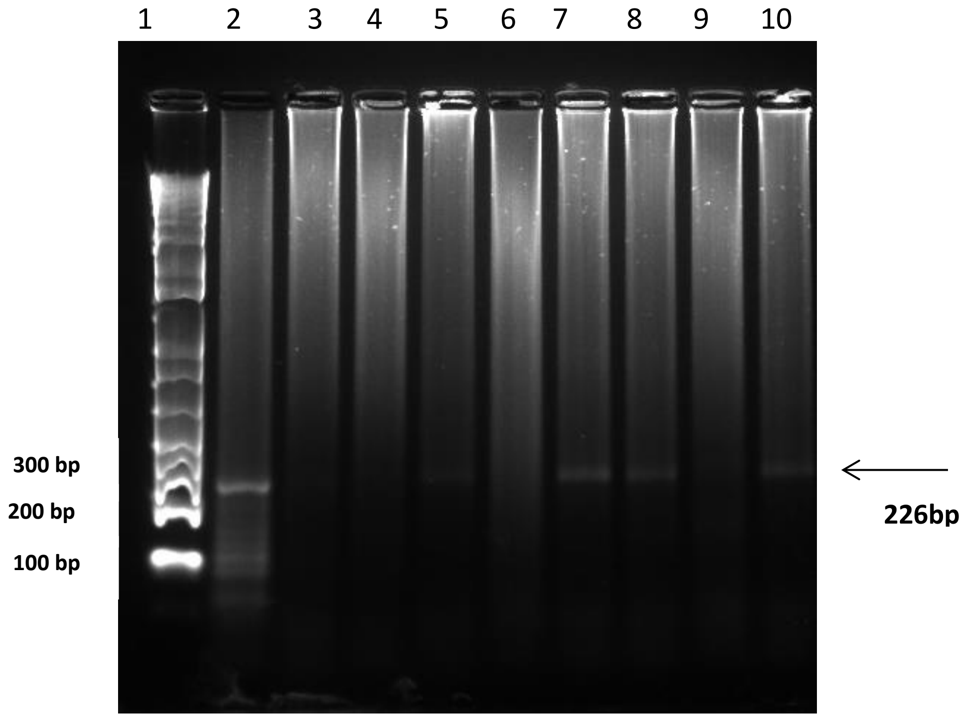


Figure 5.6 An example of PCR amplification of SAG3 locus from lung cancer patients on 1.5% agarose gel. *Lane 1* = Hyperladdder 100bp (Bioline), *Lane 2* = positive control (expected 226 bp product), *Lanes 4 and 6 are the* negative controls, *Lanes 3, 5, 7, 8, 9* and *10* are positive samples. The results were clearer on original gels by adjusting exposures and on computer version but on printed images it is difficult to see the bands.

After detecting such a high prevalence with five molecular markers we decided to use further immunohistochemical and histological methods to confirm the presence of the parasite.

5.3.2 Immunohistochemistry (IHC) detection of *T. gondii*

Immunohistochemistry was performed on 76 of the 88 tissue sections using commercial polyclonal antibodies produced in rabbits. Unfortunately we were not able to obtain the tissue section from the remaining 12 patients. All 76 sections were confirmed to be positive for *T. gondii* (Figure 5.7).

Overall the IHC showed the following results. Various numbers of cysts were detected, but surprisingly a very high number of alveolar macrophages were infected with the parasite. ICH also revealed focal groups of tachyzoites. Immunostained slides were assessed using quantitative criteria (ImageJ) to provide a percentage score which described the degree of coverage of infected tissue. For each patient (slide) three randomly selected x 400 magnification microscope fields were photographed. After the calculation of the proportional staining area to whole fields with ImageJ (as a percentage of stained pixels), the mean percentage of pixels in the stained areas was calculated for each slide. According to the final mean (intensity score) the patients were divided into three grades. Grade 1 had an intensity score below 10%, grade 2 between 10-20% and grade 3 contained all the patients with an intensity score greater than 20%. The slides were also scored for tachyzoites, infected macrophages, cysts and infected other cell types. The presence of tachyzoites, infected macrophages and other cell types represents active infection while the presence of cysts is the sign of latent infection. Example of the results for the samples and for the controls is presented in Figure 5.7. Figure 5.7A shows the positive control, RH strain tachyzoites stained with antibodies. A second positive control was lung tissue from infected wood mouse (5.7D). Two negative controls used in testing can be seen on Figures 5.5B and C. Alveolar macrophages infected with *T. gondii* can be seen in Figure 5.7F, infected other cell types in Figure 5.7G and stained cyst is visible in Figure 5.7E. In Figure 5.7H represents visible ruptured *T. gondii* cyst with free bradyzoites. Frequencies of different stages of *T. gondii* detected with IHC in the lung tissue of cancer patients are presented in Table 5.1.

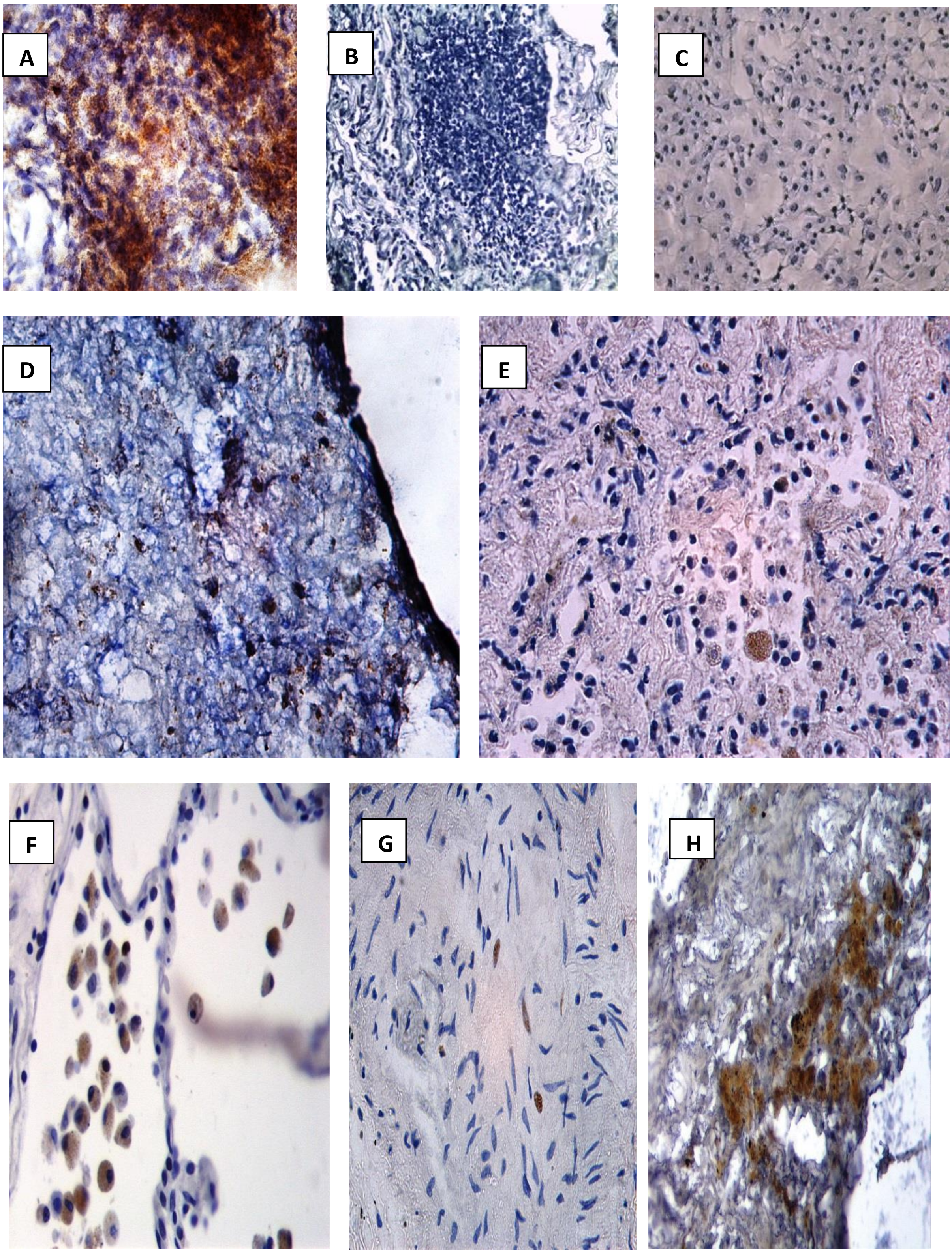


Figure 5.7: Anti-*Toxoplasma gondii* antigen immunostaining of human lung and control tissues. (A) Cell culture derived *T. gondii* RH strain tachyzoites stained with polyclonal anti –*T. gondii* antibodies. Brown staining indicates detection of *T. gondii*. (Positive control). (x400 magnification). (B) Human lung section stained with polyclonal anti –*T. gondii* antibodies with primary antibodies omitted (negative control). (x400 magnification). (C) Cells derived from a C2C12 culture (mouse myoblast cell line) which is *T. gondii* free

(negative control) and stained with polyclonal anti –*T. gondii* antibodies (x400 magnification). (D) Lung tissue from a *T. gondii* infected wood mouse (*Apodemus sylvaticus*) stained with polyclonal anti –*T. gondii* antibodies. Brown staining indicates detection of *T. gondii*. (Positive control). (x400 magnification). (E) Human lung section, from subject 1045, stained with polyclonal anti –*T. gondii* antibodies. *T. gondii* cysts can be seen. (x400 magnification). (F) Human lung section, from subject 1040, stained with polyclonal anti –*T. gondii* antibodies. Alveolar macrophages infected with *T. gondii* can be seen. (x400 magnification). (G) Human lung section, from subject 1028, stained with polyclonal anti –*T. gondii* antibodies. Fibroblasts infected with *T. gondii* can be seen. (x400 magnification). (H) Human lung section, from subject 975, stained with polyclonal anti –*T. gondii* antibodies. Ruptured *T. gondii* cysts and free *T. gondii* tachyzoites can be seen. (x400 magnification).

Table 5.1 Frequencies (%) of different stages of *T. gondii* detected with IHC in the lung tissue of cancer patients

Sample	Age	Sex	Grade	Infection stage	Cysts	Macrophages	Other infected cells types	Tachyzoites	Prevalence	Sample type
603	67	M	1	Tachyzoites	0	0	0	8.9	8.9	HEX
629	70	M	2	cyst tach macro	0.83	14.65	0	3.2	18.68	HEX
637	74	M	1	macro, tach	0	6.37	0	2.41	8.78	asthma
664	75	M	2	tach, macro, cells	0	0.77	0.46	12.1	13.33	COPD S
674	68	M	1	cyst, macro	0.38	2.5	0	0	2.88	COPD EX
757	75	M	3	macro, cells	0	20	23.5	0	43.5	COPD S
812	78	F	3	cysts, tach, cells	4.3	0	8.2	19.22	31.72	HS
813	79	M	1	other cells	0	0	1.02	0	1.02	HEX
817	74	F	3	macro, tach	0	20.45	0	7.69	28.14	HS
818	78	M	2	macro, tach	0	14.67	0	0	14.67	COPD S
819	79	M	2	macro, cells	0	16.99	2.27	0	19.26	COPD EX
821	72	M	2	cyst, tach, macro, cells	2.2	6.1	4.47	4.66	17.43	COPD S
822	67	M	2	macro, cells	0	10.71	2.1	0	12.81	COPD EX
823	84	M	1	tach, macro, cells	0	5.7	2.14	1.8	9.64	HEX
825	64	M	1	cysts, tach, cells	1.21	0	2.13	1.1	4.44	HEX
827	68	M	1	Macro	0	8.18	0	0	8.18	COPD EX
828	50	F	2	Macro	0	10.34	0	0	10.34	HS
829	79	F	3	Macro	0	34.24	0	0	34.24	HEX
832	76	M	1	tach, cells	0	0	2.98	0.33	3.31	HEX
833	69	F	3	cysts, tach, cells	0.5	0	2.78	40.69	43.97	HEX
834	69	F	1	Cyst	0.2	0	0	0	0.2	HS
835	80	F	2	Macro	0	10.39	0	0	10.39	COPD S
836	76	M	1	Tachyzoites	0	0	0	8.37	8.37	HEX
837	69	F	1	Tachyzoites	0	0	0	9.72	9.72	HNS
840	57	F	1	Cyst	2.9	0	0	0	2.9	HS
965	82	F	1	Macro	0	1.21	0	0	1.21	COPD S

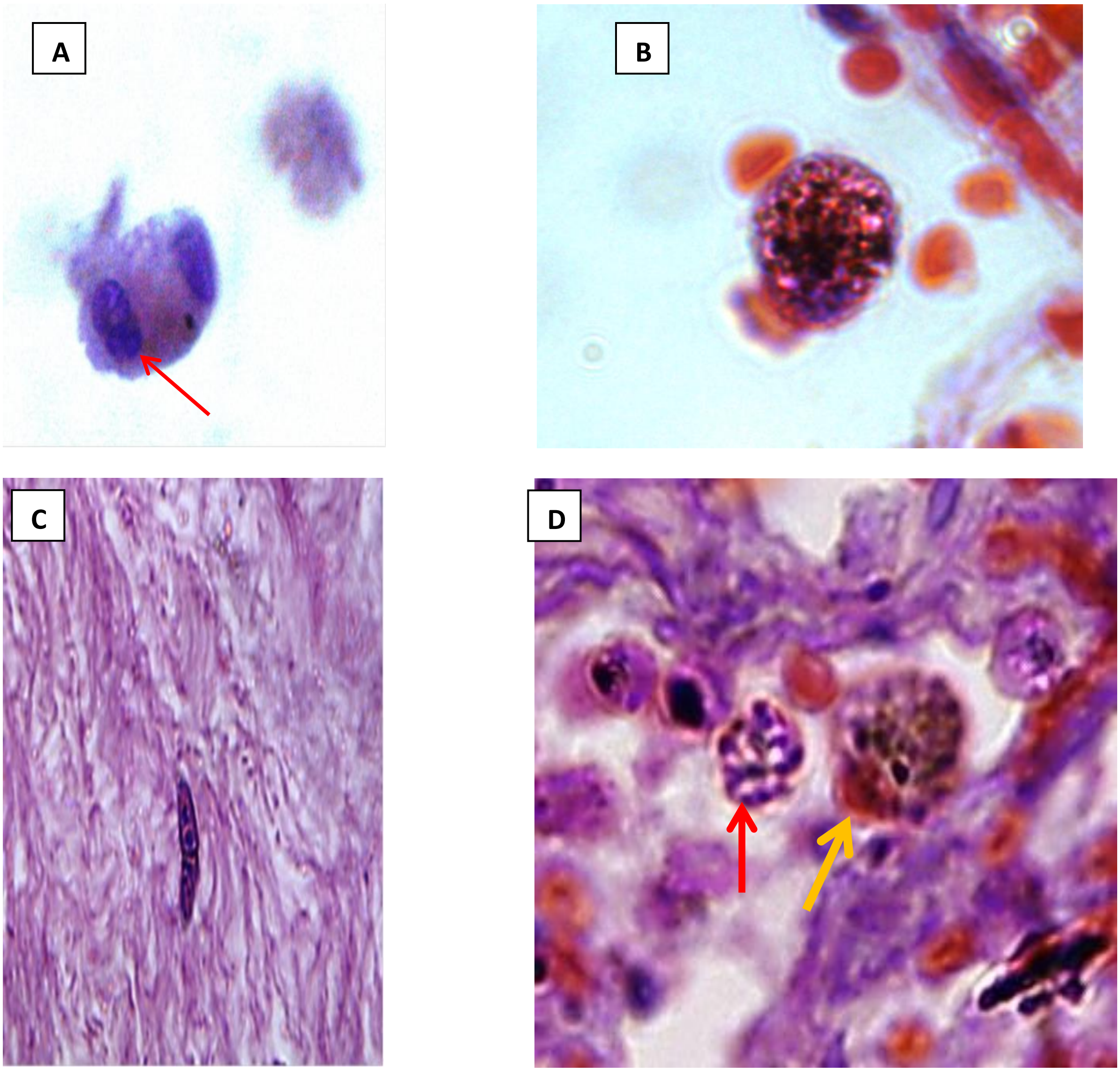
968	82	F	1	tach, cells	0	0	1.43	5.61	7.04	HEX
972	75	F	1	tach, cells	0	0	1.5	6.48	7.98	HEX
973	66	F	3	cysts, tach, cells	0.8	0	0.3	25.87	26.97	HS (COPD border)
975	71	F	2	cysts, tach	1.6	0	0	12.03	13.63	HS
976	69	F	1	Cyst	1.6	0	0	0	1.6	COPD S
979	75	F	3	cysts, tach	1.1	0	0	22.81	23.91	COPD S
985	67	M	1	cysts, tach, cells	0	0	0.3	4.81	5.11	HEX
988	73	F	2	cysts, tach, cells	0.27	0	1.9	13.14	15.31	HEX
989	71	M	3	cysts, tach, cells	1.2	0	0.4	25.1	26.7	COPD EX
997	80	M	2	cyst, macro, cells	1.2	5.07	4.3	0	10.57	COPD EX
999	60	M	3	macro, tach	0	35.07	0	8.62	43.69	COPD S
1004	67	M	3	tach, macro, cells	0	13.77	1.14	26.86	41.77	COPD S
1005	78	M	1	Macro	0	1.1	0	0	1.1	COPD EX
1006	72	M	1	Macro	0	6.45	0	0	6.45	COPD EX
1008	76	M	1	other cells	0	0	1.4	0	1.4	HEX
1010	78	M	1	macro, tach	0	4.04	0	1.11	5.15	COPD S
1013	60	M	1	other cells	0	0	1.74	0	1.74	COPD EX
1014	62	F	1	tach, macro, cells	0	2.08	0.95	0.52	3.55	COPD EX
1015	65	F	1	macro, tach	0	3.5	0	0.3	3.8	HNS
1017	70	M	1	cyst, macro	1.1	2.6	0	0	3.7	COPD EX
1018	44	F	1	macro, cells	0	5.87	0.89	0	6.76	HS
1021	66	F	1	Macro	0	8.8	0	0	8.8	HS
1025	68	F	2	macro, tach	0	10.44	0	1.08	11.52	HS
1026	77	M	1	Tachyzoites	0	0	0	0.81	0.81	COPD EX
1028	58	F	3	macro, cells	0	28.57	9.1	0	37.67	HEX
1029	66	M	2	Macro	0	11.88	0	0	11.88	COPD S
1030	61	M	2	cysts, tach	11.3	0	0	7.37	18.67	COPD EX
1032	59	M	2	tach, macro, cells	0	14.16	1.97	1.4	17.53	HS
1033	62	F	1	Macro	0	7.14	0	0	7.14	HEX

1035	63	M	2	macro, tach	0	12.43	0	0.51	12.94	HS
1036	71	F	2	Macro	0	12.37	0	0	12.37	HNS
1037	72	M	3	macro, tach	0	31.94	0	5.79	37.73	COPD S
1040	71	M	3	Macro	0	31.91	0	0	31.91	HS
1043	75	F	3	macro, cells	0	18.1	6.6	0	24.7	COPD S
1045	59	M	2	cyst, tach, macro, cells	0.59	11.78	0.9	1.69	14.96	HEX
1047	65	M	2	tach, macro, cells	0	13.53	1.57	2.32	17.42	COPD S
1049	61	F	3	macro, tach	0	34.52	0	6.12	40.64	asthma
1051	70	M	3	cysts, tach, cells	0.4	0	15.8	9.34	25.54	asthma
1052	79	M	1	cysts, tach, cells	4.2	0	0.9	1.02	6.12	COPD EX
1053	75	F	2	tach, macro, cells	0	16.63	1.88	1.34	19.85	HS
1054	67	F	3	Macro	0	20.51	0	0	20.51	HS
1060	76	F	1	other cells	0	0	0.61	0	0.61	COPD EX
1064	75	M	1	Macro	0	6.08	0	0	6.08	COPD EX
1067	52	M	2	macro, tach	0	10.91	0	1.2	12.11	HS
1068	66	M	2	macro, tach	0	15.02	0	1.75	16.77	COPD S
1069	57	F	3	tach, macro, cells	0	21.68	2.4	4.66	28.74	HEX
1070	79	M	1	macro, cells	0	0.8	0.27	0	1.07	HEX
1071	64	F	2	macro, tach	0	16.5	0	2.35	18.85	HS
1072	74	M	2	macro, tach	0	10.63	0	4.12	14.75	HEX
1101	65	F	1	cyst, tach, macro, cells	0.6	2.63	0.7	3.53	7.46	HNS

Patients' health status data presented in the last column was provided by Helen Carlin as a part of her PhD research.

5.3.3 Haematoxylin and eosin staining

All 76 tissue sections were successfully stained with haematoxylin and eosin and observed under the light microscope. The presence of structures consistent with infection by the parasite was confirmed in 67 samples. In all cases tissue cysts (Figure 5.8B, D, E, F and G) of different sizes, infected macrophages (Figure 5.7A) and other infected cell types (5.7C) were observed but no free tachyzoites could be detected. The remaining 9 samples could not be reliably confirmed as potentially infected by this method but could have possessed less visible structures such as tachyzoites.



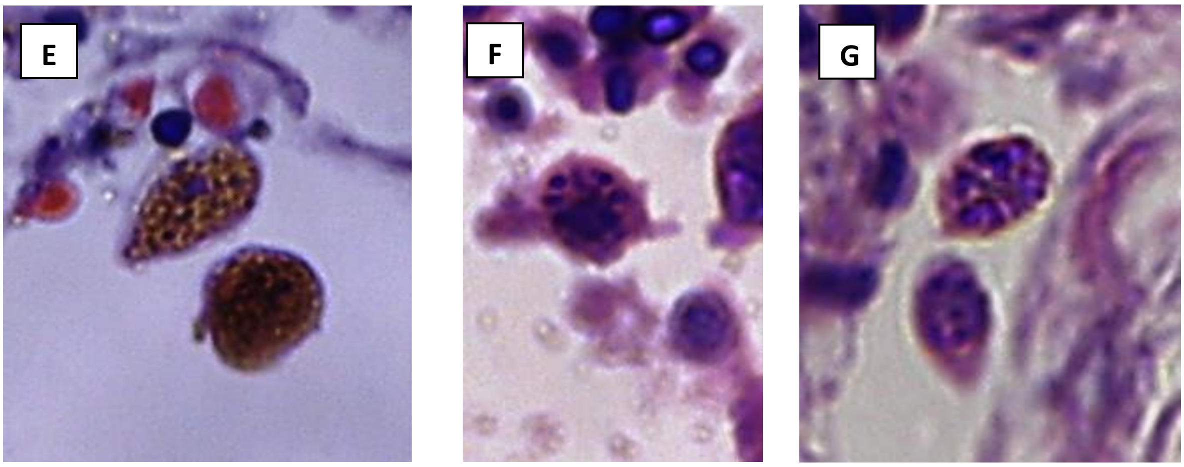


Figure 5.8: Haematoxylin and Eosin staining of human lung sections (A) Infected alveolar macrophage with four visible tachyzoites (red arrow) in subject 1040. (x400). (B) Tissue cyst from subject 1070 (x400) and (C) young tissue cyst in subject 1028. (x400). (D) Probable free bradyzoites (red arrow) observed in subject 684 possibly come from the ruptured cyst (orange arrow). Images E, F, G show different stages of tissue cysts from subbjects 821, 827 and 968 (x400).

5.3.4 Summary of prevalence

In this study we have detected 100% prevalence of *T. gondii* in a cohort of lung cancer patients using molecular and immune-histological techniques. In order to determine factors associated with increased probability of COPD/obstruction and any relationship with *T. gondii* infection, logistic regression was used. Patients were placed into 2 categories – healthy (no obstruction) or unhealthy (COPD/Asthma) and considered against factors that predicted individuals in each category. Factors included age and gender of the individual and a number of measures of smoking and *Toxoplasma* status. Average age of tested subjects was 64.64 years. *Toxoplasma* associated variables: acute infection (when infected macrophages, other infected cell types and tachyzoites were observed), latent (when only the cyst stage was detected), grade of infection, i.e. free tachyzoites (only the free tachyzoite stage was present) and *Toxoplasma* intensity (mean percentage of stained pixels for each sample). A surprisingly high number of patients (96.05%) showed evidence of an active form of infection, as defined by the presence of tachyzoites or infected alveolar macrophages (or other cell types). Only three subjects (3.95%) had the dormant cyst stage as the only stage present (Table 5.1). None of the measures (e.g. smoker status *Toxoplasma* status, gender, age etc.) were significantly associated with the probability of being healthy (i.e. no obstruction).

5.3.5 Molecular analysis and RFLP genotyping

After confirming the presence of *T. gondii* in tested tissue samples with B1, SAG1, SAG2 (5' and 3'), SAG3, haematoxylin and eosin staining and final confirmation with *T. gondii* specific antibody staining, positive samples were RFLP-genotyped with 11 markers. The genotypes were compared, identified and matched to those listed in *T. gondii* database at www.toxodb.org. Examples of successful genotyping with selected markers and RFLP-pattern diagrams from which we derived *T. gondii* genotypes (Type I, II, III or unusual types) are listed in Figure 5.9 - 5.19. From 88 tested samples all 88 were confirmed to be positive with both, $3^{/}$ and $5^{/}$ markers and all were successfully RFLP genotyped. Strain typing revealed Type I in 3 patients, Type II in 58 patients and Type III in 10 patients. Seventeen tissue samples had mixed infection of Types II and III at SAG2 loci.

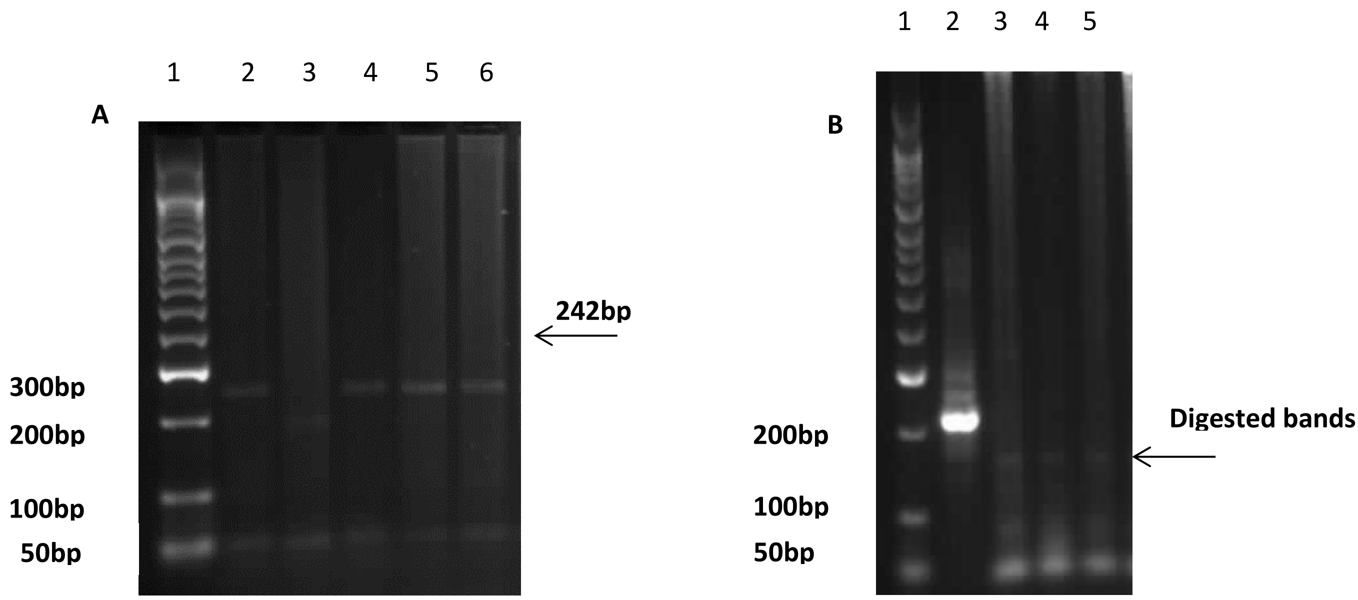


Figure 5.9 A) PCR-RFLP analysis of the SAG2 $5^{/}$ end flanking region on 2.5% agarose gel. *Lane 1* is Hyperladder II (Bioline), *Lane 2* is undigested control (242bp). *Lanes 3-6* contain samples which were incubated with *Mbo*I. This enzyme cuts Type III only while Types I and II stay undigested. The sample in Lane 3 was fully digested which indicates Type III is present. **B)** PCR-RFLP analysis of the SAG2 $3^{/}$ end flanking region on a 2.5% agarose gel. *Lane 1* is Hyperladder II (Bioline), *Lane 2* is undigested control (222bp). *Lanes 3-5* were incubated with *Hha*I which digests Type II only (present in Lane 4) while Types I and III stay undigested. All here presented bands were detected on original gels but not on reproduced prints.

All 88 tested samples were positive with SAG3 marker and all were successfully genotyped in two independent restriction digests. Type II was the most prevalent genotype and was detected in 78 cases. Type III only was found in one patient. Type I was present only as part of mixed infection with Type II in two cases and together with Types II and III in four cases. Two patients had both genotypes II and III.

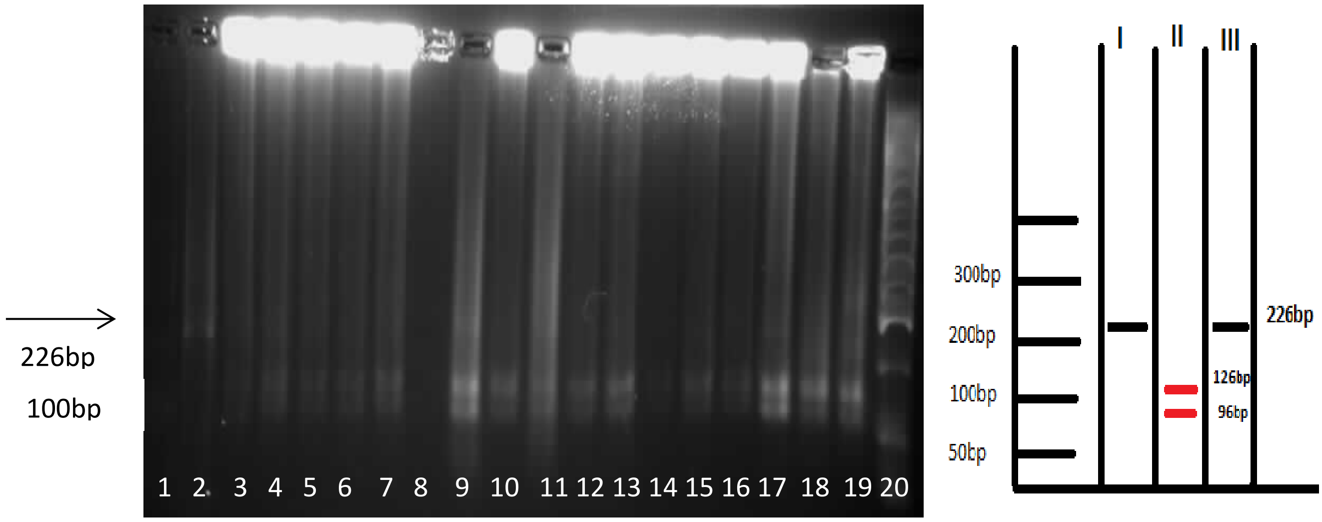


Figure 5.10 An example of agarose gel electrophoresis analysis of SAG3 PCR amplification products and restriction with *AlwN*I (cuts only type II) digests from *T. gondii* infected lung tissue samples. *Lane 2* is the undigested control, *Lane 20* contains Hyperladder 50bp (Bioline). All samples were fully digested with *AlwN*I and that means they are all Type II. The diagram shows patterns of digestion of Type II with *Alwn*I and specific restriction product sizes.

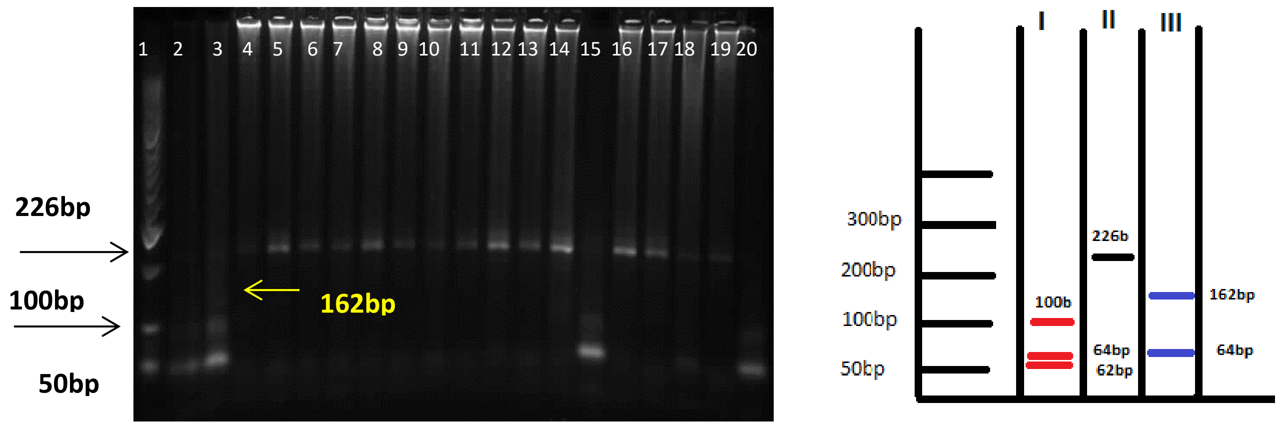


Figure 5.11 An example of agarose gel electrophoresis RFLP analysis (diagram) of SAG3 PCR amplification products and restriction *NciI* (cuts only type I and III) digests from *T. gondii* infected samples on 2.5% agarose gel with Hyperladder 50bp (Bioline) in lane 1. *Lane 4* contains undigested control. *Lanes 2, 3, 15* and *20* contain partially digested products which indicates that mixed genotypes (products from two types) are present. The samples in remaining lanes stayed undigested and that means they are all type II.

Out of 88 tested tissues samples only 43 (48.86%) were amplified and successfully genotyped with alternative SAG2 marker. RFLP detection revealed Type II in 30 samples, one sample had only Type I and no sample had only Type III. One case was a mixed infection of Types I and II and seven patients had mixed infection of Types II and III. An unusual Type I was found in four patients.

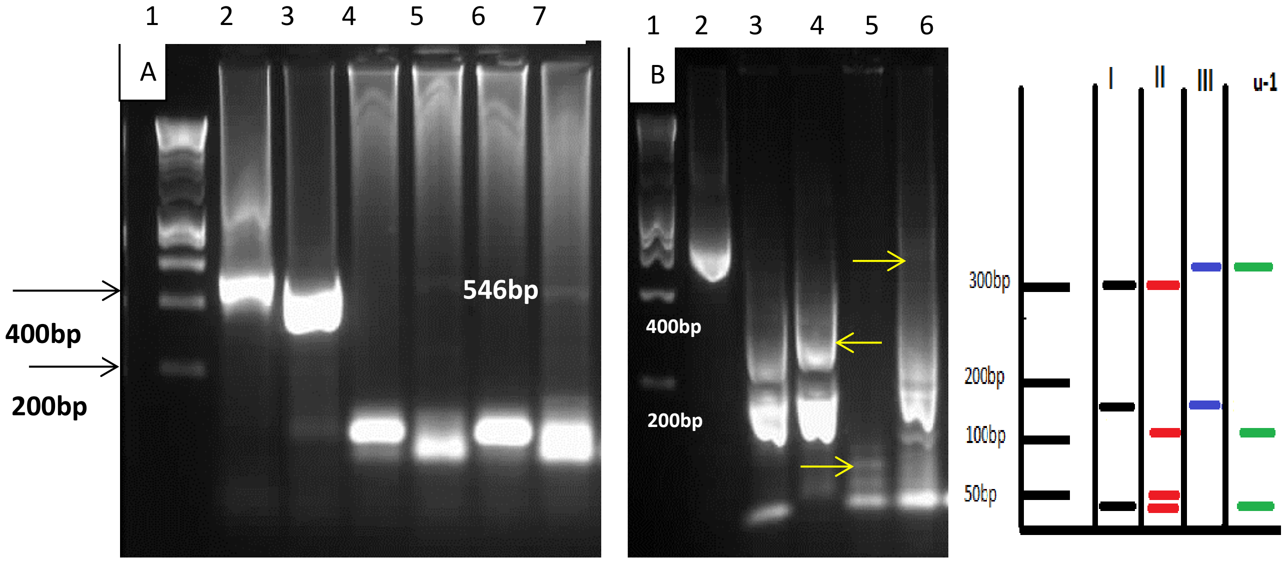


Figure 5.12 A) Agarose gel electrophoresis of alt.SAG2 marker on a 1.5% gel with 1kb Hyperladder (Bioline) in *Lane 1*. *Lane 2* is the positive control; *Lane 4* negative control and *Lanes 3*, *5* and *7* are positive samples. **B)** RFLP double digests of alt.SAG2 PCR products with enzymes *Hinf*I and *Taq*I on 2.5% agarose gel. *Lane 2* is the undigested control, digest patterns in *Lanes 3* and *4* represent type I. Lane 5 is type II but the top bands are missing. Partial digest in *Lane 6* suggests that mixed infection is present (I+II). The results were clearer on original gels by adjusting exposures and on computer version but on printed images it is difficult to see the bands (yellow arrows)

After testing tissue samples with Apico marker we found only 27 to be amplified, giving a prevalence of 30.68 %. Strain typing revealed Type I in 11 cases, followed by Type III in 9 cases and only 3 patients had Type II at this locus. In four patients we found mixed infection of genotypes II and III.

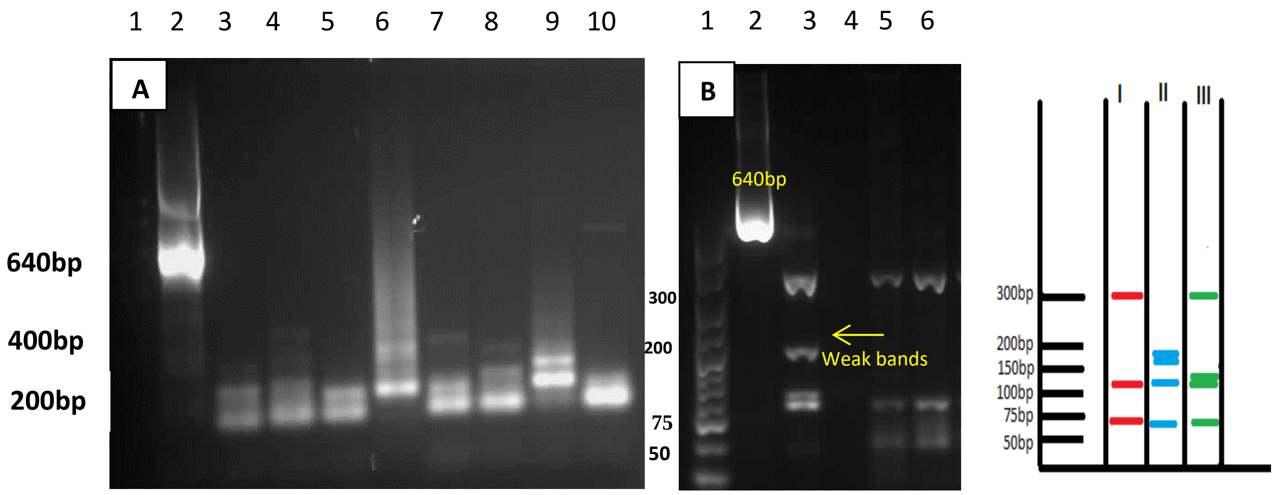


Figure 5.13 A) Apico PCR electrophoresis on a 1.5% agarose gel with Hyperladder I (Bioline) in *Lane 1*. *Lane 2* is the positive control (640bp product), *Lane 3* and 9 are negative controls. Only one positive sample is in *Lane 10*. **B)** Double digest of the PCR products with *Afl*II and *Dde*I on 3% agarose gel using Bioline Hyperladder V (Lane 1) *Lane 2* shows the undigested control, *Lane 3* represents mixed infection (partial digestion) of types II and III , in *Lanes 5* and *6* samples are both types I.

BTUB and c22-8 were the most successful RFLP markers (after SAG2 and SAG3). BTUB amplified parasite DNA in 54 out of 88 tested samples (61.36%). Interestingly with this marker Type I was the most dominant and was found in 32 cases as a single type infection, followed by Type III with 9 cases and Type II with 8 cases. Two patients had mixed infection of Type I and II and 3 patients were of mixed types II and III. With the c22-8 marker 56 out of 88 tested samples were found to be positive (63.64%). With this marker Type I one was also the most prevalent one and was found in 39 patients. Four patients had Type II infection and only one had Type III present. Mixed infections of Type I and II were found in four cases, five had mixed genotypes I and III and all three types were found in 2 patients. One unusual Type I was also detected.

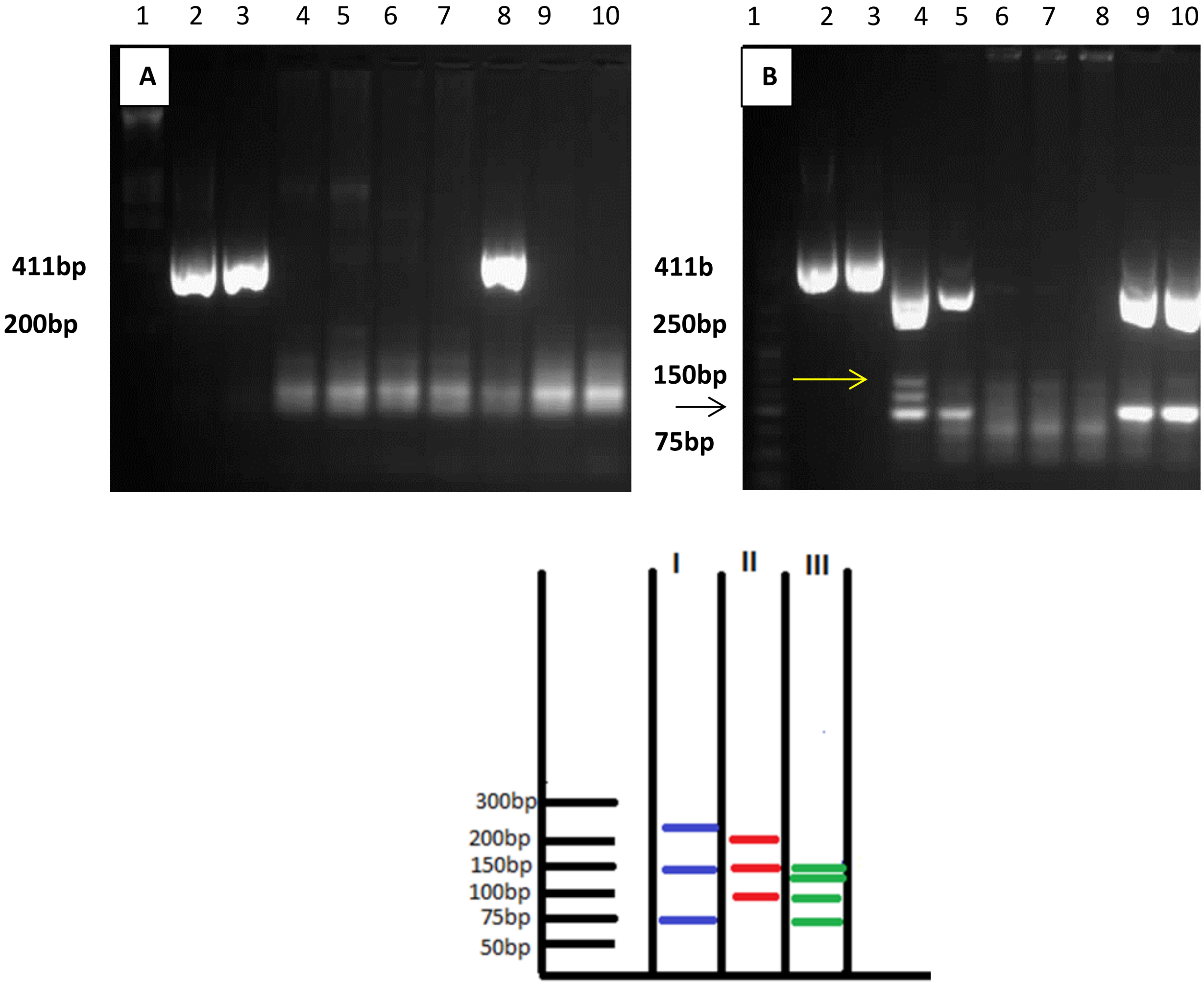


Figure 5.14 A) BTUB PCR agarose gel electrophoresis (1.5% gel). *Lane 1* is Hyperladder I (Bioline), *Lane 2* is the positive control, *Lanes 9* and *10* negative controls, *Lanes 3, 5,6* and 8 are positive samples (weak bands present). **B)** Double digest of the BTUB PCR products with enzymes *BsiE*I and *Taq*I on a 2.5% agarose gel. *Lane 1* is Hyperladder V (Bioline), *Lanes 2* and *3* are undigested controls, *Lane 4* contains mixed infection of Types III and II, *Lanes 5, 6,7,8,9* and *10* are all Type I.

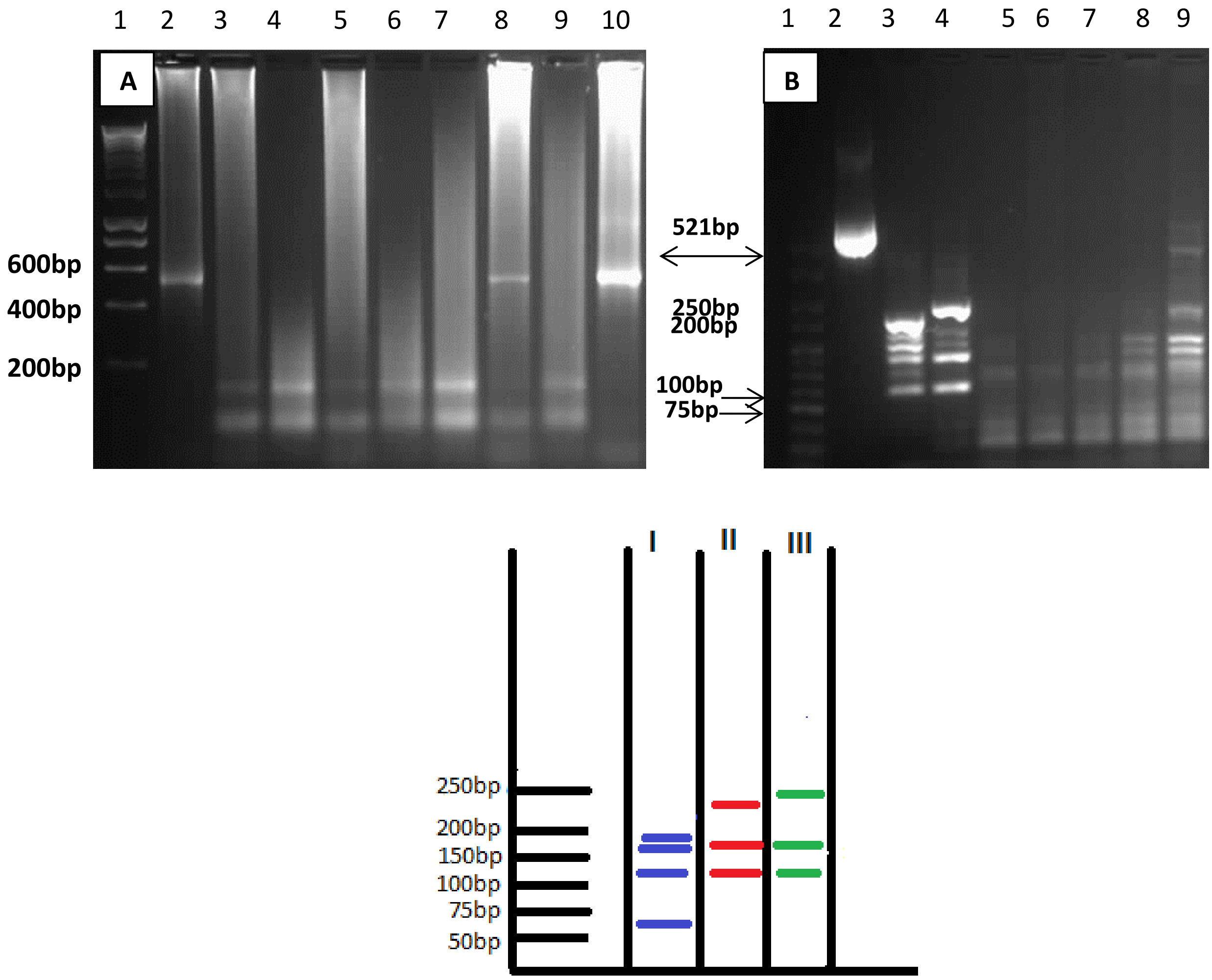


Figure 5.15 A) Nested PCR of the c22-8 marker on 1.5% agarose gel. *Lane 1* = Hyperladder I (Bioline), *Lane 2* = positive control (521bp), *Lanes 3* and *4* = negative controls, *Lanes 8* and *10* = positive samples. **B)** Double digest of the c22-8 PCR products with enzymes *BsmA*I and *Mbo*II on a 2.5% agarose gel. *Lane 1*= Hyperladder V (Bioline), *Lane 2* = undigested control, Lane 3= mixed infection (Type I and II), *Lane 4*= mixed infection (type I and III), *Lanes 5, 6, 7* = Type I, *Lanes 8* and *9* = mixed infections (Type I and III).

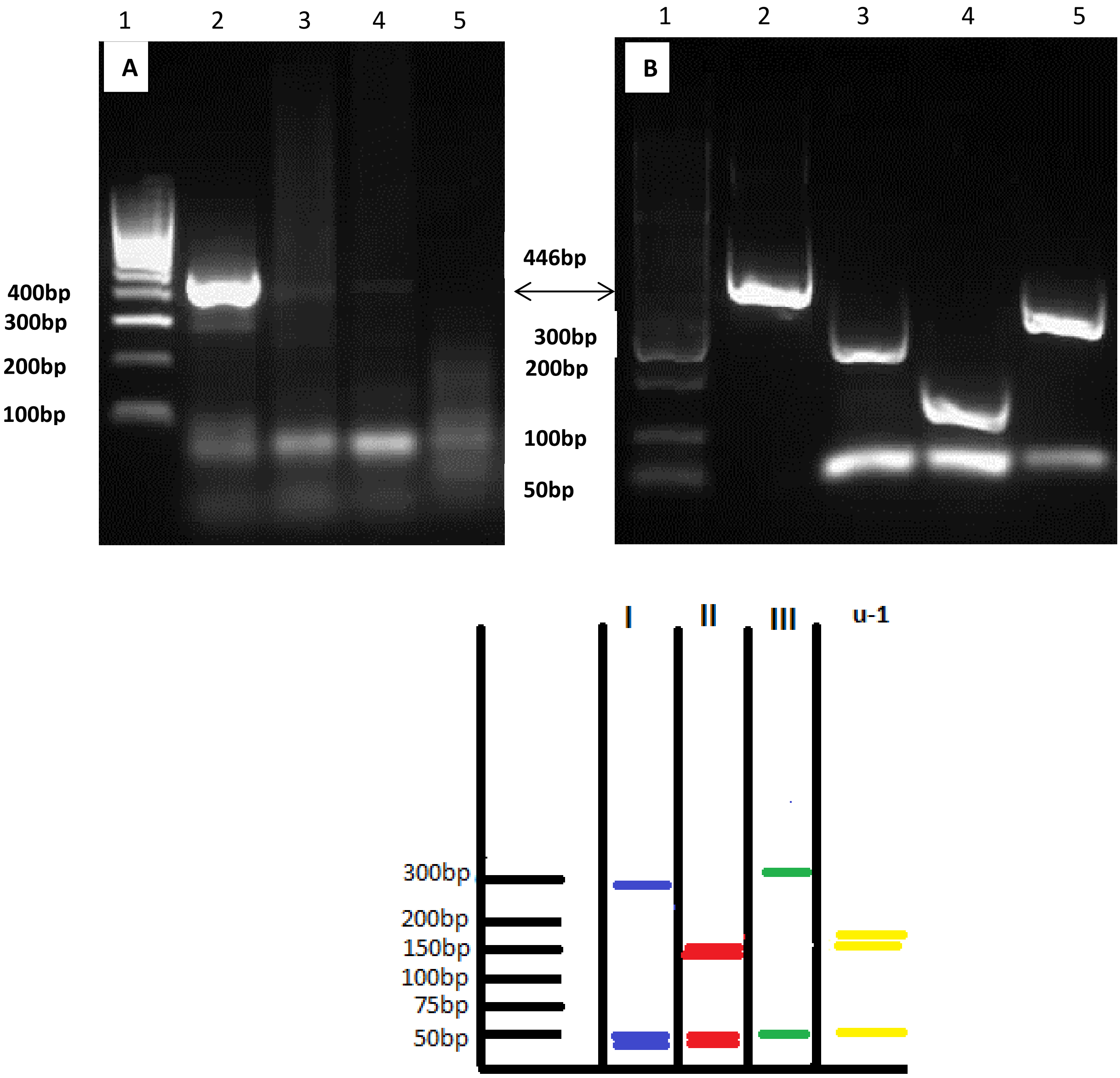


Figure 5.16 A) Gel electrophoresis of the c29-2 marker (1.5% agarose gel). *Lane 1*= Hyperladder 100bp (Bioline), *Lane 2* = positive control (446bp), *Lanes 3* and *4* = *T. gondii* positive samples, *Lane 5* = negative control. **B)** Double digest of the c29-2 PCR products with enzymes *HpyCH4*IV and *Rsa*I on a 2.5% agarose gel. *Lane 1*= Hyperladder 50bp (Bioline), *Lane 2* = undigested control, *Lane 3* = Type I, *Lane 4* = Type II, *Lane 5* = Type III.

Markers L358 and c29-2 showed similar amplification rates, with 41 (46.59%) and 42 (47.72%) samples amplified. With L358 Type I was found in 3 cases, Type II in 16 and Type III in 3. Two mixed infections of Type I and II, two of Type I and III and four of Type II and III were also found. RFLP with c29-2 revealed Type I in 12 patients, Type II in 10 and Type

III in nine patients. From ten detected mixed infections, two had Type I and II present, three had I and III and four had II and III genotypes present. Two unusual cases of Type I were also found.

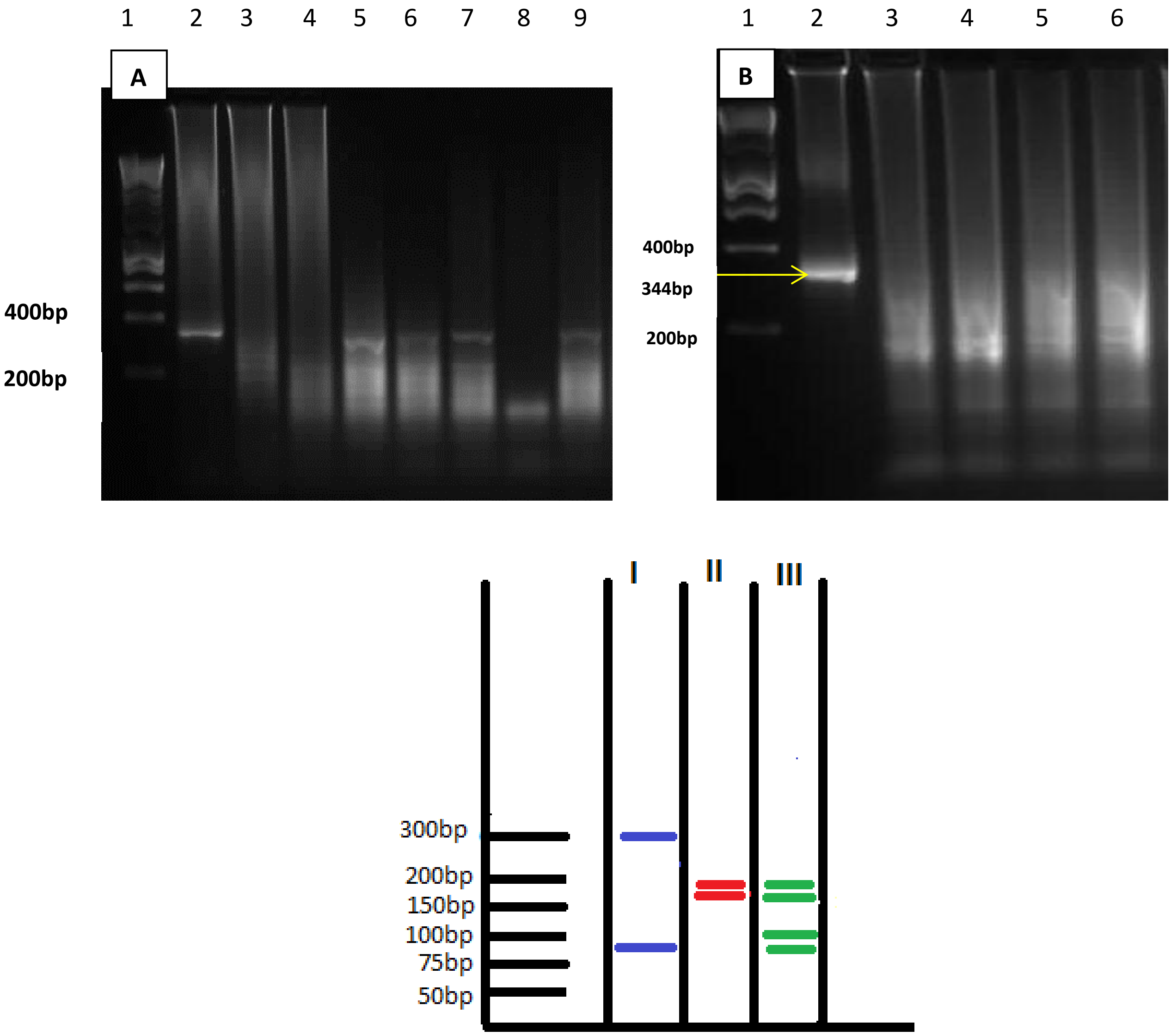


Figure 5.17 **A)** GRA6 PCR products gel electrophoresis on a 1.5% agarose gel. *Lane 1* = Hyperladder I (Bioline), *Lane 2* = positive control (344bp), *Lanes 5, 6, 7* and *9* = positive samples, *Lane 8* = negative control. **B)** Restriction digests of the GRA6 PCR products with *MseI* on 2.5% agarose gel. *Lane 1* = Hyperladder 50bp (Bioline), *Lane 2* = undigested control, *Lanes 3, 4, 5 and 6* = Types II.

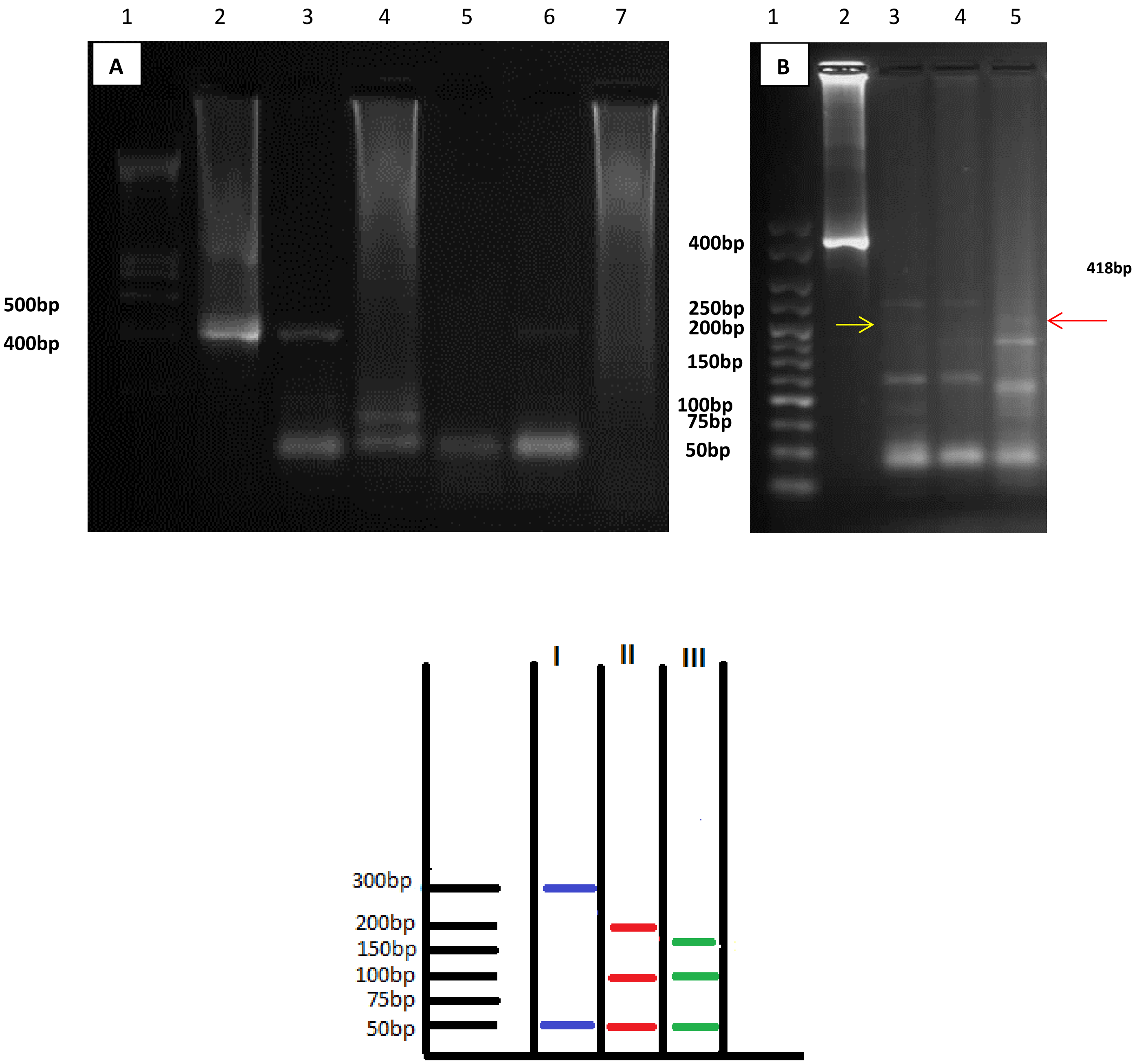


Figure 5.18 A) Gel electrophoresis of L358 marker products on a 1.5% TBE agarose gel. *Lane 1* = Hyperladder I (Bioline), *Lane 2* = positive control, *Lanes 3* and *6* = positive samples, *Lane 5* = negative control. **B)** Double digest of the L358 PCR products with enzymes *Hae*III and *Nla*III on a 2.5% agarose gel. *Lane 1* = Hyperladder V (Bioline), *Lane 2* = undigested control, *Lanes 3* and *5* = mixed infection (Type I, II and III), *Lane 4* = mixed infection (Type I and III). The results were clearer on original gels by adjusting exposures and on computer version but on printed images it is difficult to see the bands.

From all used markers PK1 had the poorest ability to amplify, when form all 88 tested samples only 9 were found to be positive (10.23%). From these, Type I was found in one case, Type II in three patients, two cases had mixed infection of Type I and III and one had genotypes II and III present. We found two unusual Types I.

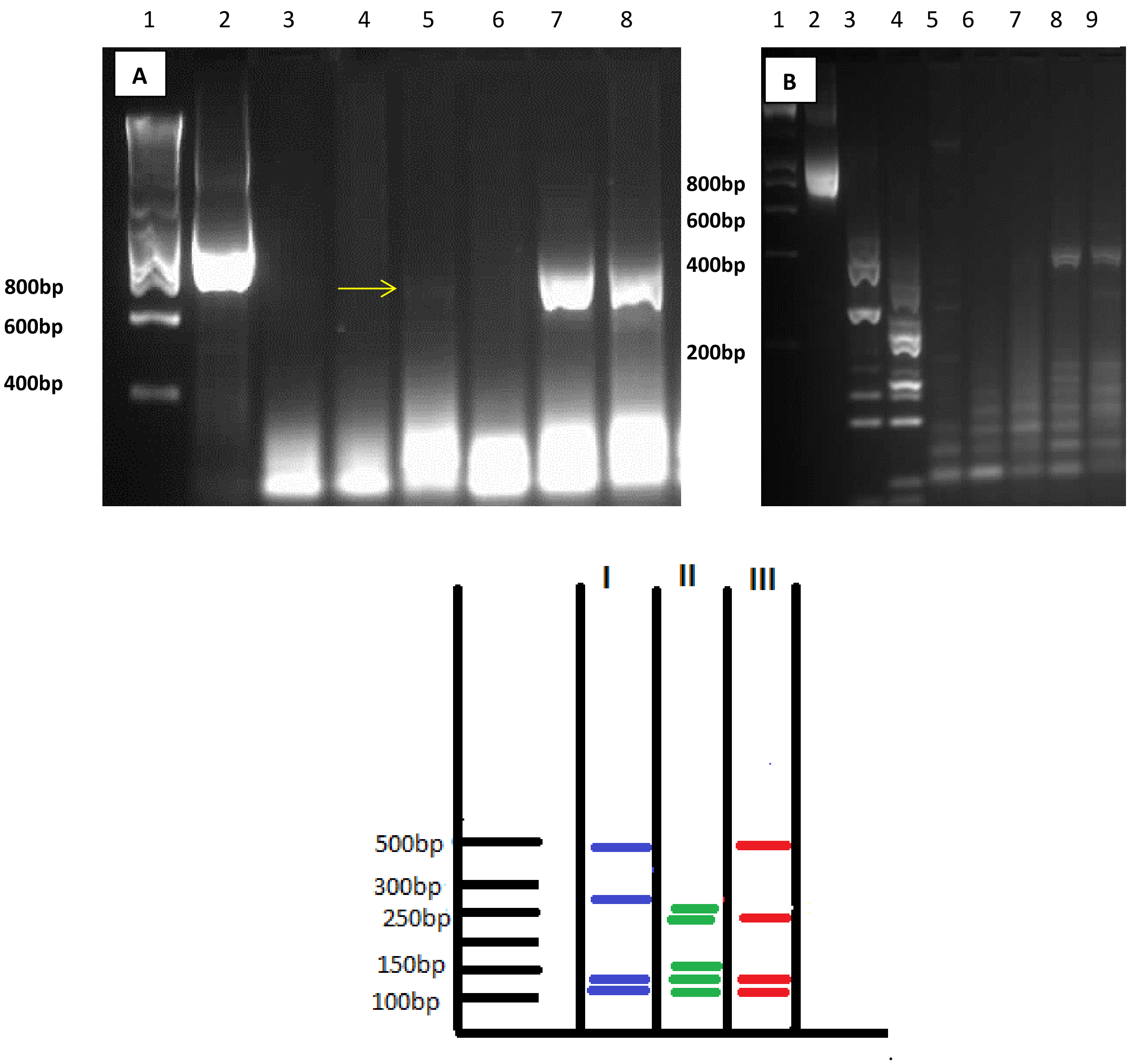


Figure 5.19 A) Gel electrophoresis of PK1 marker on 1.5% agarose gel. *Lane 1* = Hyperladder I (Bioline), *Lane 2* = positive control (903bp), *Lane 3* = negative control, *Lanes 5, 7* and *8* = positive samples. **B)** Double digests of the PK1 PCR products with enzymes *Ava*I and *Rsa*I on 2.5% agarose gel. *Lane 1* = Hyperladder I (Bioline), *Lane 2* = undigested control, *Lane 3* = mixed infection (Type I and III), *Lane 4* = Type II, *Lanes 8* and *9* = are unusual Type I

The results of successful replication for each of the individual markers are listed in Table 5.2. When comparing two markers, the one with lower detected prevalence and recorded it in the table. Using this method we created a matrix table where we compared the successful replications for all used markers. In the future this pairwise combination of two markers allows us to select combinations with the highest sensitivities (e.g. B1-SAG1, SAG2-SAG3) for fast detection of *T. gondii,* and avoid markers with low detection rates as PK1-SAG1-Su. Strain typing was performed using 11 PCR-restriction fragment length polymorphism (PCR-RFLP) markers: SAG1, SAG2, alt. SAG2, SAG3, BTUB, GRA6, c22-8, c29-2, L358, PK1, and Apico. Even though the markers used in the multilocus PCR-RFLP were developed based on sequence polymorphisms in the clonal Type I, II and III lineages (Su *et al*. 2006), non-clonal alleles (unique alleles) have already been revealed for SAG1, alt. SAG2, c22-8, c29-2, PK1 in different studies. The profiles found after digestion with restriction endonucleases were compared with the profiles of the reference strains in a virtual database, the ToxoDB database (www.toxodb.org). Recent multilocus PCR-RFLP genotyping of more than 1500 samples worldwide has revealed 221 different genotypes (www.toxodb.org). To identify the *T. gondii* strain and match it with strains listed in the database it is important to successfully genotype tested samples on all loci. If one or more loci are missing, it is not possible to match the genotype with those presented in the database. Genotyping with a single marker does not allow identification of nonclonal strains, and to determine more precisely the presence of polymorphisms in the population. It was not possible to carry out the complete genotyping in majority tested samples due to the nonamplification of some markers and the occurrence of extremely polymorphic digestion patterns. PCR-RFLP of genetic markers has been shown to be effective in the identification of *T. gondii* genotypes. However, there have been some technical problems, such as incomplete PCR amplification and restriction enzyme digestion. PCR-RFLP relies on single-copy polymorphic DNA sequences and requires a large amount of parasite DNA. The genotypes detected in lung tissue samples are described in Table 5.3 and Table 5.4. Only two samples were fully genotyped at all 10 loci. One sample (case 813) contained a single genotype only and according to the Toxodb database could be identified as a new strain. From 86 partially genotyped samples, 41 had single genotypes present and 55 were mixed infections. From the 41 single genotypes, four could be assigned or are closely matching to known T. *gondii* isolates: TgCkBr168 (database ID- 129), PRU (ID-3), isolate B73 (ID-127) and 30 could be possible new isolates. From the 55 mixed genotypes only two directly match known isolates TgCatEg57, B73 and PRU. The value of multilocus enzyme electrophoresis as a tool for genetic analysis has largely been demonstrated, but its main handicap is the need of a large quantity of purified *Toxoplasma* DNA.

Table 5.2 The matrix of successful PCR reproduction with *T. gondii* markers. Five markers (B1, SAG1, SAG2 $3^{/}$ and $5^{/}$ and SAG3 had a success rate 100% (red). Markers BTUB (green) and c22-8(orange) also showed high rates of successful replications.

Successful PCR (%)	**B1**	**SAG1**	**SAG2 5'**	**SAG2 3'**	**SAG2**	**SAG3**	**L358**	**Apico**	**C29-2**	**SAG1(Su)**	**GRA6**	**PK1**	**BTUB**	**alt.SAG2**	**c22-8**
B1	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
SAG1	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
SAG2 5'	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
SAG2 3'	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
SAG2	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
SAG3	100	100	100	100	100	100	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63
L358	46.59	46.59	46.59	46.59	46.59	46.59	46.59	30.68	47.72	23.86	46.59	10.23	46.59	46.59	46.59
Apico	30.68	30.68	30.68	30.68	30.68	30.68	30.68	30.68	30.68	23.86	30.68	10.23	30.68	30.68	30.68
C29-2	46.59	46.59	46.59	46.59	46.59	46.59	46.59	30.68	47.72	23.86	46.59	10.23	46.59	46.59	46.59
SAG1(Su)	23.86	23.86	23.86	23.86	23.86	23.86	23.86	23.86	23.86	23.86	23.86	10.23	23.86	23.86	23.86
GRA6	57.95	57.95	57.95	57.95	57.95	57.95	46.59	30.68	47.72	23.86	57.95	10.23	57.95	48.86	57.95
PK1	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23	10.23
BTUB	61.36	61.36	61.36	61.36	61.36	61.36	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	61.36
alt.SAG2	48.86	48.86	48.86	48.86	48.86	48.86	46.59	30.68	47.72	23.86	48.86	10.23	48.86	48.86	48.86
c22-8	63.63	63.63	63.63	63.63	63.63	63.63	46.59	30.68	47.72	23.86	57.95	10.23	61.36	48.86	63.63

Table 5.3 Distribution of *T. gondii* genotypes determined in lung cancer patients by PCR-RFLP. U-atypical alleles, * - SAG2 RFLP in mixed infection cannot exclude possible Type I, ** - SAG1-(Su) RFLP does not differentiate between the Types II and III

	Marker												
Genotype	**B1**	**SAG1**	**SAG2**	**SAG3**	**c22-8**	**alt SAG2**	**L358**	**Apico**	**C29-2**	**SAG1(Su)**	**GRA6**	**PK1**	**BTUB**
I	NA	NA	3	0	39	1	3	11	12	3	5	1	32
II	NA	NA	58	78	4	30	16	3	10	15**	23	3	8
III	NA	NA	10	1	1	0	13	9	9	15**	10	0	9
I+II	NA	NA	0	3	4	1	3	0	2	1	1	0	2
I+III	NA	NA	0	0	5	0	2	0	3	0	5	2	3
II+III	NA	NA	17*	2	0	7	4	4	4	0	7	1	0
I,II,III	NA	NA	0	4	2	0	0	0	0	0	1	0	0
U	NA	NA	0	0	1	4	0	0	2	2	0	2	0

Selected samples were sent for sequencing (Source Bioscience, Rochdale) to confirm the mixed infections and resolve the strain type. The fragments were sequenced with both internal primers. Samples showing mixture of two or three alleles revealed possible double peaks at known polymorphic sites indicating the presence of more alleles in the same sample (Figure 5.20-5.21). Due to the presence of strong background noise we cannot definitely confirm or exclude the presence of mixed strains.

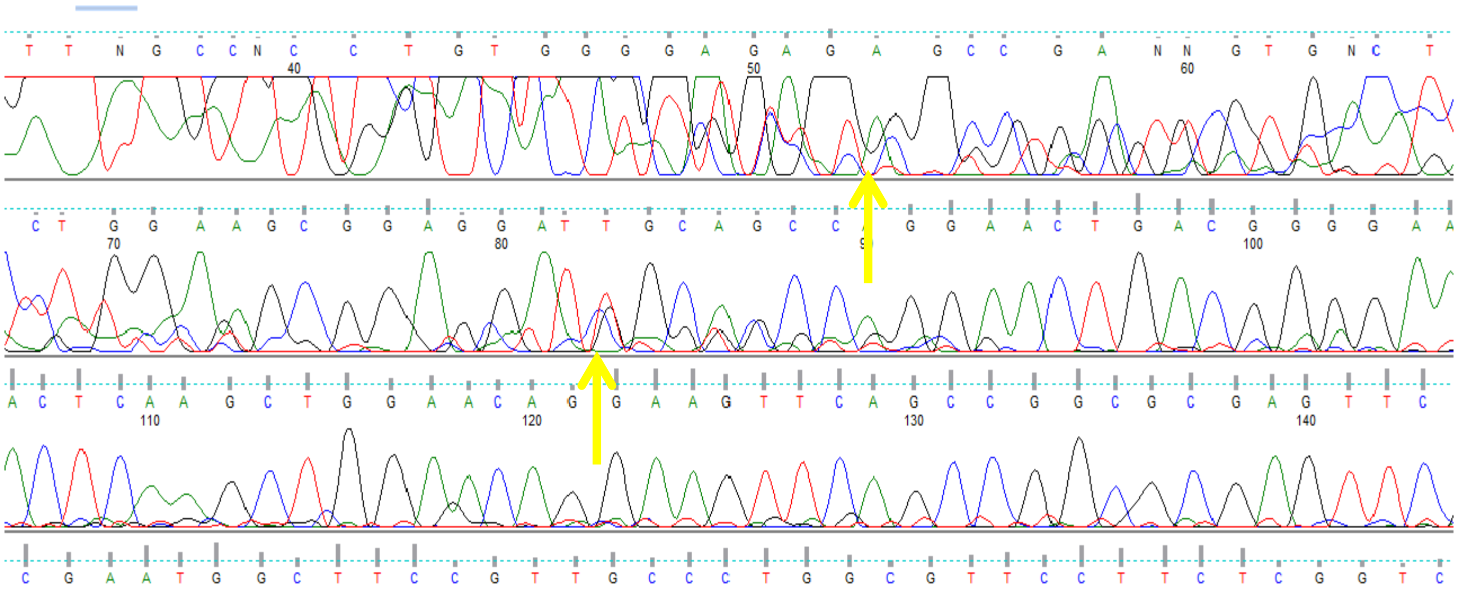


Figure 5.20 Sequence analysis of the SAG3 marker using forward primers, which reveals possible double peaks (yellow arrows) at know polymorphic sites (sample 833, contains mixed genotypes I and III).

Blast identified this sample as possible Type II strains PTG (99%), Me49 (99%), PRU (99%) and SAG3 Type I, GenBank: KT310097.1 (97%).

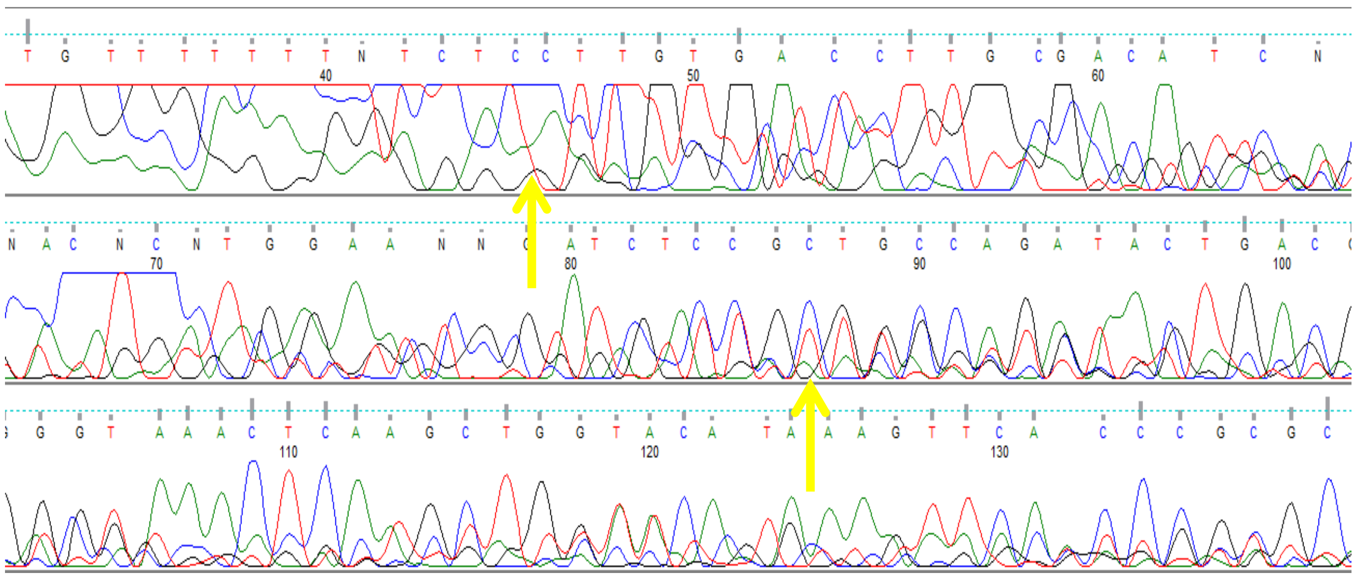


Figure 5.21 Sequence analysis of the SAG3 marker using forward primers, which reveals possible double peaks (yellow arrows) at known polymorphic sites (sample 1049, contains mixed genotypes I, II and III).

This sample was identified by BLAST as strain AF10/08U and SAG3 Type I, GenBank: KT310097.1 .Strain AF10/08U was detected in Tunisia (Boughattas *et al*. 2010) and was associated with congenital toxoplasmosis. In this study this strain had mixed alleles II and III on SAG3 loci.

Table 5.4 *T. gondii* genotypes determined in 88 human patients with lung cancer

Sample	SAG2 5'	SAG2 3'	SAG2	SAG3	L358	Apico	C29-2	SAG1(Su)	GRA6	PK1	BTUB	alt.SAG2	c22-8	Sample type
603	I or II	II	II	II	N	I	N	N	N	N	II	N	N	P
625	I or II	II	II	II	N	II	N	N	N	N	N	N	N	P
629	I or II	II	II	II	N	III	N	N	N	N	I	N	II	p
637	I or II	II	II	II	N	I	N	N	N	N	I	N	N	asthma
663	I or II	II	II	II	N	I	N	N	N	N	N	N	N	COPD S
664	I or II	II	II	II	N	N	N	N	N	N	I+II	N	N	COPD S
674	I or II	II	II	II	N	III	N	N	II	N	II	N	N	COPD EX
684	I or II	II	II	II	N	I	N	N	N	N	N	N	N	HS
757	I or II	II	II	II	N	N	N	N	N	N	N	N	N	COPD S
812	I or II	II	II	II	N	II	III	N	II	N	N	II	I	HS
813	I or II	II	II	II	III	I	III	II or III	II	I	I	u1	I	HEX
817	III + mix	II + mix	II+III	III	III	N	N	N	I,II,III	N	III	N	I	HS
818	I or II	II	II	II	II+III	I	III	N	II+III	N	I	II	I	COPD S
819	I or II	II	II	II	III	N	N	N	N	N	I	II	I	COPD EX
821	III + mix	II + mix	II+III	II	N	N	N	N	III	N	N	N	I + III	COPD S
822	I or II	II	II	II	III	N	III	N	III	N	N	II+III	I	COPD EX
823	I or II	II	II	II	N	N	II	N	II	N	III	II	I	HEX
824	I or II	II	II	II	II+III	N	II	N	II	N	II	N	N	COPD S
825	I or II	II	II	II	N	I	II+III	N	II+III	N	I	I+II	I	p
827	I or II	II	II	II	N	N	u1	N	II	N	III	N	I	COPD EX
828	III + mix	II + mix	II+III	II	II	III	II+III	N	II+III	N	N	II+III	I	HS
829	I or II	II	II	II	I+II	I	III	N	III	N	N	N	I	HEX
832	III + mix	II + mix	II+III	I,II,III	II	N	I	N	II+III	N	I	II	I	HEX
833	III + mix	II + mix	II+III	I+II	N	I	II	N	N	N	N	N	I	HEX
834	I or II	II	II	I,II,III	III	N	II	I	II+III	N	N	I	I	HS
835	I or II	II	II	II	N	N	I	N	III	N	I	u1	I	COPD S

836	I or II	II	II	II	I + III	N	III	N	III	N	N	II	I	HEX
837	I or II	II	II	II	N	III	II	II or III	II	N	N	II	I	HNS
839	I or II	II	II	I,II,III	I+II	N	II+III	N	I + III	N	N	II+III	I	HEX
840	I or II	II	II	II	N	III	I	N	I	N	N	II	I	HS
918	I or II	II	II	II	II	I	I	N	III	N	II	II	II	COPD S
963	I or II	II	II	II	N	N	N	N	I + III	N	II	N	N	HS
965	III + mix	II + mix	II+III	II	I	N	II	N	I + III	N	N	N	N	COPD S
968	III	I or III	III	II	II	N	III	N	II	N	III	II	III	HEX
972	I or II	II	II	II	III	N	I+II	N	I+II	N	III	N	I	HEX
973	I or II	II	II	II	N	N	II	N	III	II	I	II	I	p
975	I or II	II	II	II	N	N	I	N	I	N	N	II	I	HS
976	I or II	II	II	II	N	N	I	I	II	N	I	II	I	COPD S
979	III + mix	II + mix	II+III	I+II	N	N	I+II	II or III	II	N	I+II	II	I+II	COPD S
985	III + mix	II + mix	II+III	II	N	N	II+III	N	II+III	N	I	II+III	I+II	HEX
988	I or II	II	II	II	N	N	N	II or III	I	N	I	N	I	HEX
989	I or II	II	II	II	N	N	N	I +(II or III)	II	I + III	I + III	II	I+II	COPD EX
997	I or II	II	II	II	II	N	II	II or III	II	N	I	II	I	COPD EX
999	I or II	II	II	II	N	N	I	N	I	N	N	N	N	COPD S
1004	I or II	II	II	II	I	N	u1	N	III	N	III	II	I	p
1005	III + mix	II + mix	II+III	II	II	N	I	I	I + III	I + III	III	II	I + III	COPD EX
1006	I or II	II	II	II	II+III	N	III	N	II	N	N	N	I	COPD EX
1008	III + mix	II + mix	II+III	II	III	I	I	II or III	II	II+III	II	II+III	I+II	HEX
1010	II	II	II	II	N	N	I + III	N	N	N	I + III	N	N	p
1013	I or II	II	II	II+III	II	III	I	N	II	N	III	N	I	COPD EX
1014	III + mix	II + mix	II+III	II	I	N	I + III	N	N	N	I + III	N	II	COPD EX
1015	III + mix	II + mix	II+III	II	N	N	N	N	N	N	I	N	N	p
1016	I or II	II	II	II+III	II	N	N	N	II	N	N	N	N	HEX
1017	III	I or III	III	II	N	N	I + III	N	N	u1	II	II	I + III	COPD EX

1018	I or II	II	II	II	N	II+III	I	N	N	N	N	N	N	HS
1020	I or II	II	II	II	N	II+III	N	u1	II+III	N	I	II	I	p
1021	I or II	II	II	II	I+II	II	N	N	N	N	I	N	N	HS
1023	I or II	II	II	II	I + III	N	N	N	N	N	N	N	N	HNS
1025	III	I or III	III	II	N	II+III	N	II or III	N	N	N	N	N	HS
1026	I or II	II	II	II	N	N	N	N	III	N	I	N	I	COPD EX
1027	I or II	II	II	II	N	N	N	N	II	N	I	N	I	HNS
1028	I or II	II	II	II	II	N	N	N	N	N	III	N	N	HEX
1029	I or II	II	II	II	N	N	N	N	I + III	N	I	N	I	COPD S
1030	I or II	II	II	II	N	N	N	N	II	N	N	u1	N	COPD EX
1032	III + mix	II + mix	II+III	II	II+III	N	N	N	N	N	N	N	u2	HS
1033	III + mix	II + mix	II+III	II	N	N	N	N	I	N	N	II	I,II,III	HEX
1035	I or II	II	II	II	II	N	I	N	N	N	I	II+III	I	HS
1036	III + mix	II + mix	II+III	II	N	II+III	N	II or III	N	N	N	N	N	HNS
1037	I or II	II	II	II	N		N	II or III	N	N	N	N	II	COPD S
1040	I or II	II	II	II	N	III	N	N	N	N	I	u1	I	HS
1043	I or II	II	II	II	II	N	III	N	N	N	I	N	N	COPD S
1045	I or II	II	II	II	N	N	II	N	II	N	I	II	N	p
1047	III	I or III	III	II	N	N	N	II or III	II	N	N	II	I	COPD S
1049	I or II	II	II	I,II,III	II	N	N	N	N	N	N	N	I	asthma
1051	I or II	II	II	II	II	N	N	N	N	II	N	N	I + III	asthma
1052	I or II	II	II	II	N	N	N	II or III	N	N	N	II	N	COPD EX
1053	I or II	I or III	I	II	N	III	N	II or III	II	N	I	II	I + III	HS
1054	I or II	I or III	I	II	II	N	N	N	N	N	I	N	N	HS
1056	III	I or III	III	II	II	N	N	II or III	II	II	I	II	N	HS
1060	I or II	I or III	I	I+II	II	N	N	N	N	N	I	II	I	COPD EX
1064	III	I or III	III	II	III	N	N	N	N	N	N	N	N	COPD EX
1067	III + mix	II + mix	II+III	II	III	N	N	II or III	N	N	I	II	I	HS
1068	III	I or III	III	II	III	N	N	N	N	N	N	N	N	COPD S

1069	III	I or III	III	II	N	N	N	N	N	N	I	N	N	HEX
1070	III	I or III	III	II	III	III	N	u1	N	u1	I	II	I	p
1071	III + mix	II + mix	II+III	II	III	N	N	II or III	II	N	II	II+III	N	HS
1072	III	I or III	III	II	N	N	II	N	III	N	I	II	I,II,III	p
1101	I or II	II	II	II	III	N	N	N	N	N	I	N	N	HNS

N= negative sample (yellow), HS = Healthy smoker, HEX = Healthy ex-smoker, HNS = Healthy non-smoker, COPDS = patient with Chronic Obstructive Pulmonary Disease), COPD EX= Ex COPD patient, p= pilot study, no data on health status are available. *T. gondii* genotype Type I-pink colour, Type II –blue, Type III-green, Type II or III type- brown, Type I or II-dark orange, Type I or III-purple, all three Types present –red.

5.3.5 Scanning Electron Microscopy and Transmission Electron Microscopy

Selected tissue samples were sent to Salford Analytical Services (SAS) at the University of Salford for the analysis by Scanning electron microscopy (SEM) and Transmission electron microscopy (TEM). Structures similar to *T. gondii* stages (Figure 5.22) were observed using both types of microscopy.

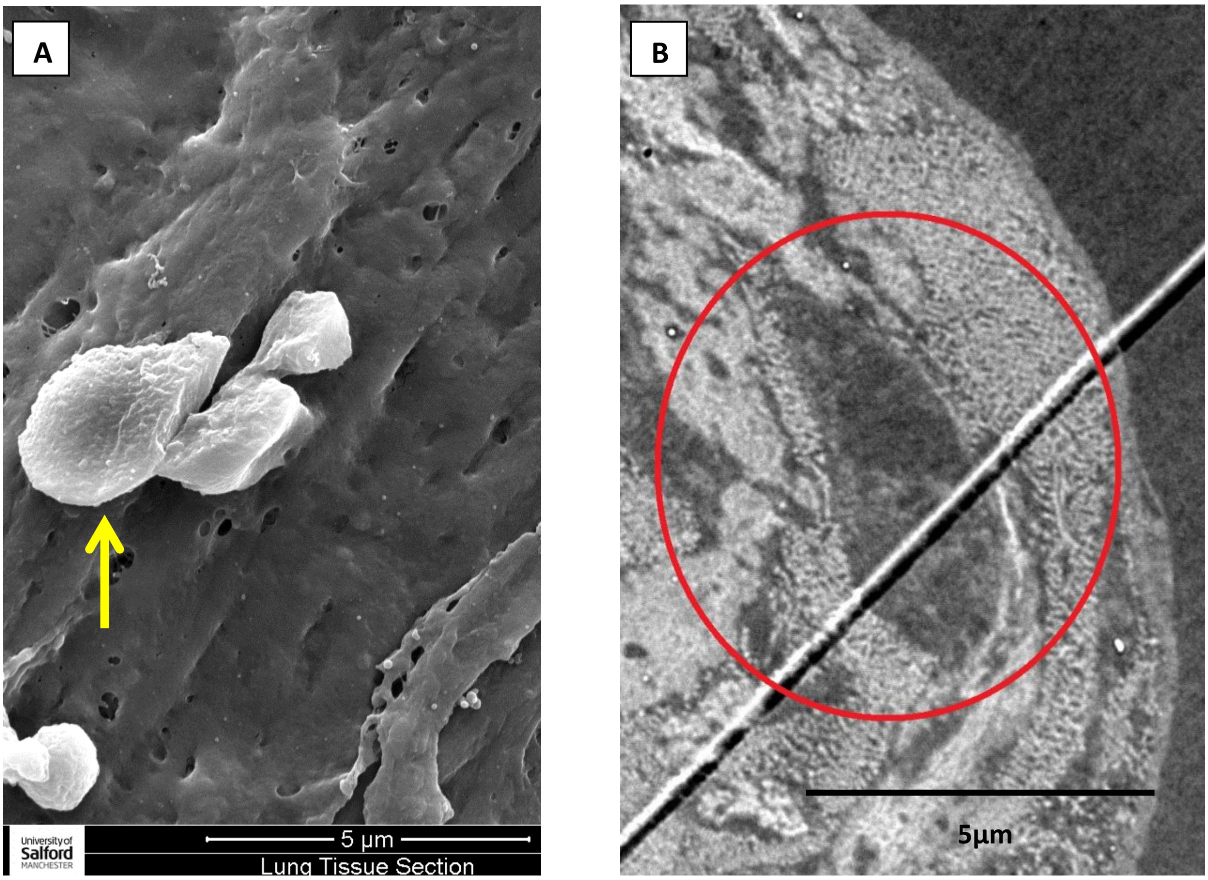


Figure 5.22 A) SEM of the lung tissue sample from case 979. A tachyzoite like structure (yellow arrow) can be observed. **B)** TEM of the macrophage from the patient 1040. A tachyzoite like structure can be observed in the red circle.

5.4 Discussion

In this study we investigated the prevalence of *T. gondii* in clinical samples from patients with lung cancer. We used four commonly used PCR markers (SAG1, SAG2, SAG3 and B1), histological and immunohistochemical staining to confirm the presence of the parasite. The prevalence ranged from 23.86% (SAG1Su) up to 100% (SAG1, SAG2, SAG3 and B1). Haematoxylin and eosin staining confirmed parasite like structures in all tested tissue sections and the staining with polyclonal *T. gondii* antibodies detected the parasite in all 76 stained sections. This study has revealed 100% prevalence in this cohort of lung cancer patients. These observed prevalences are unusually high but correspond with others studies, as the prevalence is estimated to be between 15 % up to 85% depending on geographical location (Dubey and Beattie, 1988). Around 50% of the adult population in Germany is infected and as many as 90% of adults in Paris are seropositive (Hökelek *et al.* 2013). The most recent studies in the UK show infection in the range of 7 to 34% (Joynson *et al.* 1992; Pappas *et al.* 2009; Flatt *et* al 2013; Public Health Wales 2014). Recent reports show that latent toxoplasmosis can possibly play an important role in manifestation of many psychiatric and neurological disorders such as bipolar disorder, personality disorder, Parkinson disease (Miman *et al.* 2010), Alzheimer disease (Kusbeci *et al.* 2011), OCD, epilepsy (Yazar *et al.* 2003), schizophrenia (Torrey and Yolken, 2005; Niebuhr *et al.* 2007), autism and brain tumours (Thomas *et al.* 2012; Vittecoq *et al.* 2012) but not much is known about the prevalences in cancer patients. In the study performed by Yuan (Yuan *et al.* 2007), the positivity rates of T. *gondii* IgG in nasopharyngeal carcinoma (46.15%) and rectal cancer (63.64%) groups were significantly higher than the other cancer groups, pulmonary carcinoma (4.55%), breast cancer (9.53%), gastric carcinoma (10.00%), hepatocellular carcinoma (14.29%) and uterine cervix carcinoma (12.50%)., demonstrating that there is a likely to be a stronger association between *T. gondii* infections and some kinds of cancer. Thomas and Lafferty (2011) in their study compared national figures from 37 countries and found that brain cancers are more common in countries where *T.gondii* infections are more common. Brain cancer was 1.8 times more common in countries where *T.gondii* was most common, than in those where it was virtually absent. The high percentage in the here presented study might be due to the fact that all patients have cancer and are immunocompromised which increases their susceptibility to *T. gondii*. Cancer is associated with defects in cell-mediated immunity and chemotherapy predisposes to the development of toxoplasmosis.

Selected samples were sequenced to resolve strain type whenever there were unclear enzyme digestion results or possible mixed genotypes present. In these cases, sequence analysis revealed possible double peaks at known polymorphic sites, indicating the presence of multiple alleles but due to the presence of background noise this cannot be fully confirmed. However, even sequence analysis has a limited capacity to detect mixtures where one component is in significant excess over another. As both, restriction analysis and sequencing independently confirmed the presence mixed strains in the PCR products, we are confident that this was not an artefact.

In this study a combination of different independent markers was used. Most of these markers had good sensitivity and were able to detect *T. gondii* DNA in a small amount of tissue. This technique was suitable for direct genotyping but it was not possible to genotype all tested samples on all loci. A high number of mixed infections were detected, with some samples having all three genotypes present. Mixed infections have been previously reported in England and Wales (Aspinall *et al.* 2003). This was explained by possible ingestion of more than one type of parasite in food products containing meat originating from multiple animals (Aspinall *et al.* 2002). This also suggests that intermediate hosts are more often infected with more than one strain as previously thought. The majority of tested subjects were of higher age so that congenital transmission is unlikely to be the main source of infection, but it cannot be fully excluded. Congenital transmission is the result of primary infection acquired by an immunologically naïve patient during pregnancy (Ajzenberg *et al.* 2002) and so this route of transmission does not explain the high number of mixed infections. Congenital infection acquired during the early stages of pregnancy also often results in severe foetal deformation or death. Infection acquired later may result in milder symptoms or neonate may be asymptomatic at birth. It is possible that the patients were infected with mixed oocysts or tissue cysts from contaminated food during their life, or with combination of all three routes. The information on whether these patients ever had blood transfusion or underwent organ transplantation is not available, as this could be another possible source of infection. Only three patients had SAG2 Type I infection. This genotype has been linked to congenital transmission in humans (Howe *et al.* 1995) and ocular toxoplasmosis (Grigg *et al.* 2001). Type I was unexpectedly predominant with markers c22-8 and BTUB. However the majority of tested patients had SAG2 Type II genotype present which has been reported as most prevalent genotype within both humans and animal populations (Owen *et al.* 1999). This genotype produces a high cyst burden in mice and is prone to reactivation in experimentally immunocompromised mice (Howe *at al.* 1995). This corresponds with the high number of cysts detected in tested cohort of present study. Only ten patients had SAG2 Type III infection. These strains have been found to be common in animals, but are reported to be associated with human cases of toxoplasmosis significantly less often (Howe *et al.* 1995). Overall, 55 mixed infections were detected. In general, there is no clear correlation between strain genotype and symptomatology. In immunocompromised patients, the three SAG2 genotypes are related to neurological diseases, with pneumonia or other related lung diseases (Fuentes *et al.* 2001). There is no data available on other markers, and possible links to particular diseases, as all original genotyping was based on the SAG2 marker only. The results presented here demonstrate that is possible to use direct genotyping with standard *T. gondii* RFLP markers on clinical tissue samples, and if full characterization is not always obtained, at least partial characterization is available. The nested PCR is fast and highly sensitive and the use of direct genotyping also avoids misleading results due to artificial selection in relation to the culture process. The use of cell culture or mice to isolate and grow the parasite from the clinical samples might produce sensitive variation in the observed genotyping frequencies due to an effect of differential selection of the strains (Frenkel *et al.* 1997). Extreme caution needs to be applied when claiming a correlation between *T. gondii* genotypes and toxoplasmosis clinical presentation until there is a more complete knowledge regarding the parasite genotypes circulating in the human population.

Pathological diagnosis of acute toxoplasmosis is most reliably established by demonstration of tachyzoites. With haematoxylin and eosin staining, tachyzoites can be seen in alveolar lining cells, filling macrophages, or free in the alveolar space. The surrounding inflammation is usually mononuclear, but mixed inflammation, vasculitis, necrotizing granulomas, abscesses and giant cells can be observed. Immunohistochemical staining with avidin-biotin peroxidase increases the sensitivity of routine staining. The usefulness of PCR for the diagnosis of the pulmonary toxoplasmosis remains unestablished. PCR would detect the tissue cyst present in dormant infection as well as the tachyzoites in acute infection. The diagnosis is established by the observation of tachyzoites in broncho-alveolar lavage fluid or transbronchial biopsy specimens. Reports have shown that the lung is frequently involved in fatal cases of pulmonary toxoplasmosis, where 55% of the patients died, even when pneumonia was not diagnosed on clinical grounds (Pomeroy *et al*. 1992). Therapy for acute toxoplasmosis is indicated in immunocompetent patients with significant organ dysfunction and in all immunocompromised patients.

Many bacterial and protozoan pathogens have clonal population structures and specific clones are often associated with particular disease (Tibayrenc *et al.* 1990). Araujo in his study observed that different strains of *T. gondii* induced different cytokine responses in CBA/Ca mice (Araujo *et al.* 2003). In previous reports the majority of human toxoplasmosis cases were associated with infection by the Type II genotype, often associated with toxoplasmosis in AIDS patients with some fatal consequences (Fuentes *et al.* 2001). The Type I strain in humans has been linked to a high frequency of ocular toxoplasmosis and congenital toxoplasmosis (Fuentes *et al.* 2001). Type I and atypical strains have never been reported in asymptomatic or benign congenital toxoplasmosis (Ajzenberg *et al.* 2002). Type III is not very common in human population, but Grigg (2001) observed atypical recombinant I and III genotypes in patients with ocular toxoplasmosis and AIDS. Carme suggested that is possible that atypical/recombinant genotypes are associated with severe toxoplasmosis at least in immunocompetent patients (Carme *et al.* 2002). Type III has been often associated with African countries, but as we do not know the background and the origin of the patients, we cannot exclude they are come from this area.

All our patients have been diagnosed with lung cancer and subsequently underwent treatment and probably chemotherapy. Many of them had high number of infected alveolar macrophages, which is a characteristic of Type II infections. Some of the patients suffer from asthma and COPD and are on medication, which supresses the immune system. All these patients are at risk of reactivation of latent infection, leading to toxoplasmic encephalitis or pulmonary toxoplasmosis; which can complicate and delay their treatment or lead to death. We suggest that all cancer patients should be monitored for *T. gondii* to prevent complications during their treatment.

CHAPTER 6:

General Discussion

The protozoan intracellular pathogen *Toxoplasma gondii* has a worldwide distribution and can infect all warm-blooded vertebrates, causing severe disease in immunocompromised humans. Felids are the definitive hosts and other mammals can serve as intermediate hosts that can be infected by ingested oocysts, tissue cysts or congenitally (Dubey *et al.* 2001). It is estimated that about 30% of the human population is infected but some areas can exceed 40% (Pappas *et al.* 2009). Little is known of *T. gondii* genotypes and prevalence in wild natural populations of rodents and humans in the UK. In this study we developed methodologies that can be used to investigate these questions. We have investigated the prevalence and genotypes of *T. gondii* in a wild rodent population of *Apodemus sylvaticus* in an area relatively free of cats and in a cohort of human cancer patients. We have used *T. gondii* standard markers and novel molecular tools used in population genetics to investigate the rodent population, and human samples were also tested with histological and immunohistochemical methods to verify the reliability of PCR approaches.

6.1. Prevalence of *T. gondii* in natural populations of wood mouse

In recent years there has been an increased interaction between humans, domestic and wild animals that could result in a greater exchange of pathogens. There is a close relationship between domestic and sylvatic life cycles, therefore contact between herds and wild animals can give rise to economic losses relating to reproductive problems caused by *T. gondii.* Rodents are often considered as relevant marker species to assess environmental contamination by *T. gondii,* as felids prey on them and so the prevalence in carnivores is expected to be higher. It has been experimentally confirmed that rodents infected with *T. gondii* lose their innate fear of cat odours (Vyas *et al.* 2013) and this can increase the chance of the transmission to the definitive host. In total, brains from 126 wood mice from Malham Tarn Nature Reserve were tested with five *T. gondii* markers (B1, SAG1, SAG2, SAG3 and GRA6) and 34.92% (95% CI: 27.14% - 43.59%) of all mice were infected. A total of 24/76 (31.58%, 95% CI: 22.19%-42.74%) of male and 20/50 (40%, 95% CI: 27.59%-53.84%) of female mice were found to be positive for *T. gondii.* No significant difference was found in prevalence in males and females ($\chi 2$ =0.863, D.F. = 1, P = 0.353). A total of 17 juveniles, 65 young adults and 44 adults were present in this cohort of 126 of which 7 (41.18%), 27 (41.54%) and 10 (22.72%) were PCR positive for *T. gondii.* There was no significant age prevalence effect detected (P = 0.23). However prevalence of infection is known to accumulate with age in some species and so larger species with longer life expectancy are more likely to show this trend than smaller species, such as mice. Additionally small rodents can be more susceptible to infection and die soon after contracting the infection and this may explain the lack of an age prevalence effect usually seen. Four different genetic clusters of mice were detected in this population and interestingly there was a significant difference between each genetic population in *T. gondii* infection prevalence ($\chi 2$ =7.950, D.F. = 2, P = 0.018). Twenty two families of wood mice were identified in this collection in the study site with an average *T. gondii* prevalence in families of 32.55% but with no significant differences in individual families ($\chi 2$ = 0.284, D.F. = 1, P = 0.5943). Our data showed that the prevalence of *T. gondii* in natural mammal populations is high and the parasite can be perpetuated in the absence of felids possibly via the congenital route. High levels of vertical transmission have been documented in *Mus musculus*, *M. domesticus* and *A. sylvaticus* (Owen and Trees, 1998; Marshall *et al.* 2004). The majority of epidemiological studies in wild animals have been based on serological diagnostic methods (Franti *et al.* 1976; Jackson, 1986; Webster, 1994; Dubey *et al.* 1995; Smith and Frenkel, 1995; Hejlicek *et al.* 1997; Jeon and Yong, 2000; Jakubek *et al.* 2001; Morsy *et al.* 2001; Yin, 2010; Mercier, 2013; Elamin, 2014; in comparison the PCR–based studies have shown higher prevalence when used in wild mice populations with prevalences ranging from 10.4% (Vujanic *et al.* 2011), 13.6% (Kijlstra *et al.* 2008), 23.9% (Yan *et al.* 2014), 29% (Zhang *et al.* 2004) up to 40.78% (Thomasson *et al.* 2011) and 59% (Marshall *et* al. 2004). Unfortunately we were not able to obtain the blood and do serological tests on sampled mice and were therefore unable to directly evaluate the sensitivity of PCR over serology. The sensitivity of different assays can influence detection of infection and so makes it difficult to compare the prevalences among studies using different methods. For example Araujo *et al.* (2010) isolated *T. gondii* from a rat which was identified as seronegative using MAT. Direct methods (PCR) can also be negative in seropositive individuals (Iqbal *et al.* 2007). Due to a dormant infection the antibody prevalence can be also related to age of the host due to the accumulation of exposure to parasite with age.

6.2. *Toxoplasma* genotypes circulating in wildlife

In Europe and North America three main clonal lineages (types I, II and III) are predominantly found (Howe and Sibley, 1995), with type II being more frequent in Europe. Recently a fourth clonal type (haplogroup 12) has been discovered in North America (Khan *et al.* 2011). A few isolates originating from remote area of South America have been analyzed and showed different multilocus genotypes suggesting that other lineages could circulate in other parts of the world (Darde *et al.* 1998). Su *et al.* (2003) constructed a phylogenetic tree of the Apicomplexa using rRNA sequences and showed that these three clonal lines formed a single branch with two other apicomplexa (*Hammondia, Neospora*) which had a common ancestor around 12 million years ago. Limited genetic diversity between and within clonal lineages indicates that they have evolved recently from a common ancestor, 10 000 years ago at the most (Su *et al.* 2003). Separate co-existing domestic and sylvatic cycles of *T. gondii* transmission may explain the distribution of archetypal and atypical genotypes in different hosts. Strain genotype also influences host immune response. The strain virulence is not the same according to the host, for example type I strains are highly virulent in mice but are not pathogenic in rats (Zenner *et al.* 1999; Li *et al.* 2012). Information on *T. gondii* strains circulating within the UK is limited. To determine the strains from the Malham Tarn Nature Reserve all 44 positive mice were RFLP genotyped at SAG2, SAG3 and GRA6 loci. In our study, only five mice had only Type II genotype on all three loci and the remaining 39 positive mice had mixed infections (23 had the combination of Type II and Type III alleles and 16 had all three genotypes present). This corresponds with genotypes observed in a previous study (Thomasson *et al.* 2011), when all three clonal lineages were detected in this locality. Type II was the most prevalent, but type I was more frequent than expected as this type is rare in wild life in Europe. Our study using three RFLP markers did not identify any atypical strains, but possibly adding more markers at other loci could identify novel genotypes in this population. Mixed infections in intermediate hosts have been reported previously (Ajzenberg *et al.* 2002, Aspinall *et al.* 2003, Dubey *et al.* 2003, Pena *et al.* 2006). Detection of mixed infections is of particular interest in epidemiological studies. Experimental studies showed that non-canonical genotypes can develop when a cat ingests prey infected with *T. gondii* of more than clonal type followed by a sexual cross but this is believed to be rare in nature (Sibley *et al.* 1992). Genotypes different from clonal type II were found in Spain, Poland, Portugal and Germany in chickens, pigs and pigeons (de Sousa *et al.* 2006; Waap *et al.* 2008; Berger-Schoch *et al.* 2011).

6.3. Material and methods improvements

Another problem was the aspect of sampling. Most previous genotype data are based on studies from humans and domestic animals, but the few isolates originating from other animals (arctic foxes, sea otters and some others) revealed identification of atypical multilocus genotypes or some unusual alleles. Emerging new clonal types may show different levels of virulence in intermediate hosts. In summary our results show no distinctive genotype dispersion pattern for the study area, but indicate some differences in the prevalence of types among wild rodents. For future study it would be beneficial to work with a larger sample size, this would be especially useful for genetic studies and analyses of the rodent family structures and *T. gondii* infection prevalence. As the life span of mice is only a few months and they breed about three times per year, the sampling period over two years caused some complications when we tried to assign the individuals into populations and families. We observed higher infection frequencies from the autumn samplings and this can be possibly linked to the source of infection and transmission. The mice should be sampled at least twice per month for the maximum period of two years. Ten microsatellite mice markers were suitable for the population studies but it would be useful to use some *T. gondii* microsatellites to analyse the circulating parasite strains. Maybe testing DNA extracted from the blood could be performed to detect early infection before the parasite has created tissue cysts. In a Brazilian study DNA was extracted from the buffy coat instead of the whole blood and the detected prevalence was 48.6% (Kompalic-Cristo *et al.* 2007). Some serological and histological studies on different tissue types (lungs, brain, and liver) should be compared with molecular results. Isolation and culturing of the parasite would be a big advantage for the parasite genetic analysis and should be employed if possible. This however is very cumbersome for large epidemiological studies.

6.4. Human toxoplasmosis

Humans may acquire *T. gondii* infection by ingesting food or water containing sporulated oocysts, by consuming raw or undercooked meat or congenitally. Major outbreaks of toxoplasmosis in Canada and Brazil were food- and waterborne related (Bowie *et al.* 1997; de Moura *et al.* 2006). The high seroprevalence rates found among vegetarians (Hall *et al.* 1999) in India support the importance of environmental or water contamination. In humans, primary infection is usually asymptomatic or does not require medical treatment. If the symptoms occur, they are usually mild influenza like symptoms, rarely hepatosplenomegaly and lymphadenopathy can occur, and are often self-limiting (Krick and Remington, 1978). Latent infections are known to induce morphological and behavioural changes in infected persons (Webster *et al.* 2001). In immunosuppressed and immunodeficient patients *T. gondii* infection can have fatal consequences. The lungs represent the second most common site of *T. gondii* infection in people with AIDS (Marche *et al.* 1989). Those individuals with defects in cell mediated immunity represent a large group of the population at risk of pulmonary toxoplasmosis.

6.5. *Toxoplasma gondii* in cancer patients

Most of the studies that investigate the link between cancer and *T. gondii* infection are based on serological prevalence detection of the parasite in cohorts of cancer patients rather than by direct investigation of tissue samples. In our study we tested 88 samples from patients with lung cancer and other lung diseases (COPD, asthma) for the presence of *T. gondii* using molecular and immune-histological and a prevalence of 100% (95% CI: 95.19%-100%) was detected. This is a surprisingly high prevalence when compared to some previous studies. Overall 8.38% of examined patients with malignant neoplasms in China were positive for antibodies against *T. gondii*. However, when nested PCR detection was used on the same sample, only 3.55 % of these patients were positive (Wang *et al.* 2015). In another study from China (Cong *et al.* 2015) much higher seroprevalence was detected, with 35.56% of the cancer patients overall being positive for anti-*T.gondii* IgG. The highest prevalence of infection (Table 6.1), in this study, was observed in lung cancer patients (60.94%) followed by cervical cancer patients (50%). Among 356 cancer patients, 21 (5.9%) cases were found to be IgG-positive and 8 (2.3%) were IgM-positive, and five of them were found to have both IgG and IgM antibodies (Shen *et al.* 2014). The total seroprevalence of *Toxoplasma* infection among these patients was 6.8%. A study in Iran concluded that 45.2% of cancer patients were seropositive for *T. gondii* (Ghasemian *et al.* 2007). High seropositivity rates were detected in women with breast cancer (86.4%).

Table 6.1 Seroprevalence of *T. gondii* detected among different cancer patients groups (Wang *et al.* 2015, Yuan *et al.* 2007).

Cancer type	*T. gondii* prevalence (%)
Lung cancer	60.94
Cervical cancer	50.00
Breast cancer	60.94
Nasopharyngeal carcinoma	46.15
Rectal cancer	63.64
Gastric carcinoma	10.00
Hepatocellular carcinoma	14.29

The problem is that serological studies lack both sensitivity and specificity. In few published studies IgM remained positive in patients with *T. gondii* infections for 12-18 months and the IgG avidity remained low for a maximum of 4 months. In the United States, less than 5% of people with AIDS and toxoplasmosis produce IgM antibody to *T. gondii*, and the rise in levels of IgG antibody is distinctly unusual (Mills *et al.* 1986). A variety of serological tests such as the dye test (DT), modified agglutination test (MAT), enzyme-linked Immunoabsorbent assays (ELISA), Immunoabsorbent agglutination assay (ISAGA), indirect fluorescent antibody test (IFAT) and indirect haemaglutination assays (IHA) are used to detect different antibody classes or antigens.

6.6. Pulmonary toxoplasmosis

Since pulmonary pneumonia in immunocompromised patients is most often due to reactivation of *T. gondii* infection, these patients generally have serological evidence of past infection (Caterall *et al.* 1986). Pomeroy and his team (1992) observed that pulmonary toxoplasmosis appeared in two basic settings, 33% of reported cases occurred in patients who had no underlying illness linked with immunosuppression and 67% of the patients were immunocompromised. Pulmonary toxoplasmosis only rarely develops in immunocompetent patients (Weller *et al.* 1988) with only 9 cases documented of *T. gondii* pneumonia. The symptoms are often nonspecific; cough is non-productive until there is coinfection with other microbes. The most common findings on chest X-Ray is the presence of diffuse interstitial infiltrates but these feature are nonspecific. How often pulmonary toxoplasmosis is clinically manifest is unknown, but the condition may be undiagnosed. Definitive diagnosis of toxoplasmosis requires either histological demonstration of the organism or isolation by culture.

Toxoplasmosis has been frequently described to be associated with some specific malignancies such as acute and chronic leukaemia, lymphoma or multiple myeloma (Yazar *et al.* 2004). In addition, patients being treated with antineoplastic drugs for tumours such as breast, ovary and lungs have been associated with toxoplasmosis (Yazar *et al.* 2004). While the medical significance of a high prevalence of infection with *T. gondii* in lung cancer patients is clear, the biological relationships between parasite infection and lung cancer are not. In laboratory animal models, the mouse and rat, it has been shown that *T. gondii* infection is related to the balance between levels of inducible Nitric Oxide Synthase (iNOS) and Arginase-1 (Arg-1) expression (Li *et al.* 2012): high iNOS, low Arg-1 being associated with resistance and the reverse associated with sensitivity. Interestingly, peritoneal macrophages from inbred lines of rats are *T. gondii* resistant (high iNOS/low Arg1), while alveolar macrophages from the same lines are sensitive to infection and express low iNOS and high Arg-1 expression (Zhao *et al.* 2013). Thus it appears, at least in rats that epigenetic differences in gene expression define alveolar cells as having a different degree of susceptibility to *T. gondii* infection. In studies, in laboratory animal models, infection with *T. gondii* causes impaired anion secretion in airway tissues by inhibiting the P2Y2 receptor (Guo *et al.* 2015). This could interfere with lung function by affecting the correct maintenance of airway surface liquid as occurs in cystic fibrosis. Furthermore, studies in laboratory rodents have shown that glucocorticoids, drugs commonly used in cancer patients, stimulate *T. gondii* growth in cells both *in vitro* and *in vivo* (Wang *et al.* 2014). Little is known as to whether parallels exist in humans, but these studies on model systems raise questions about how we view the relationship between lung cancer and *T. gondii* infection. Recent reviews claim that human macrophages have iNOS activity, albeit induced by other stimuli rather than that inducing nitrite production in murine macrophages and so may be derived from other cellular sources (Fang *et al.* 2004; Schneemann *et al.* 1997). The high frequency of *T. gondii* infection in our lung cancer patients raises questions about whether the two conditions are linked. It is unlikely that there is a direct cause and effect linkage as there are no reported causative effects of *T. gondii* infection on producing cancers, as far as we are aware. However, many types of cancer can cause immunomodulatory effects on affected tissues and individuals (Franklin *et al.* 2014; Zhao *et al.* 2015) and immunosuppression is a known risk factor associated with toxoplasmosis (Canessa *et al.* 1992, Montoya *et al.* 2004 Ajzenberg *et al.* 2009; Ahmadpour *et al.* 2014). Thirugnanam *et al.* (2013) suggested that *T. gondii* infection can promote initiation and progression of cancer by modifying the miRNAome in brain cells. Interestingly there is some scientific evidence indicating that parasitic infections induce antitumor activity against certain types of cancers. It is possible that asymptomatic *T. gondii* infection stimulates the immune system and may have anti-tumour activity. The inhibitory effect of *T. gondii* on cancer growth has also been shown in cell culture and a mouse model (Darani *et al.* 2012). Suppression of neoplastic growth has been detected for some bacteria, protozoa and helminths (Gambashidze *et al.* 2012) and that certain parasitic compounds have are able to inhibit cancer growth (Darani *et al.* 2009). In study conducted by Miyahara *et al.* (1992) injections of rats with *T. gondii* lysate antigen resulted in a significant reduction in tumour formation. In another study significant reduction in tumour size was observed in mice injected with *T. gondii* antigen absorbed in alum adjuvant (Darani *et al.* 2009). Also inhibition of Lewis lung carcinoma by *T. gondii* has been reported through induction of Th1 immune responses and IL-12 and interferon production (Kim *et al.* 2007). In a study conducted by Seyedeh *et al.* (2015) the frequency of low titer antibody against *T. gondii* in cancer patients was significantly higher than the frequency of low titer antibody against this parasite in healthy people and he concluded that exposure to this parasite may be related to resistance to cancer. In this study the age of the patients was below 70 years and all were in early stages of the different types of cancer. *Toxoplasma* tachyzoites fixed in formalin in' study Suzuki *et al.* (1986) were shown to induce antitumor activity against EL4 lymphoma in C57B1/6 mice. Also it is possible that metabolic products of *T. gondii* have inhibitory effects on cancer cells (Choo *et al.* 2005).

In our study all tested cancer patients were *T. gondii* positive and some of them had mixed strain type infection. We do not know how different strains and genotypes interact with the immune system, for example whether they can protect against the cancer or infection could be associated with manifestation of the disease, and what their impact might be on the health of the tested subjects. Further studies with a wide range of tested animal species, cell cultures and different strains needs to be done to gain a deeper understanding of the mechanism of *T. gondii* infection and dissemination. We suggest performing serological, histological and molecular testing on a larger group of patients and to compare the data with a control group of healthy patients. Different types of tissue from patients with different types of cancer should be tested and patients from the youngest age group available should be included to avoid the age prevalence effect. Molecular detection of parasite DNA in peripheral blood is of the greatest significance in immunocompromised patients (Costa *et al.* 2009, Daval *et al.* 2010). Vujanic *et* al (2012) observed a decline in parasitaemia in a patient with reactivated toxoplasmosis during anti-parasite treatment using real-time PCR. If possible, living parasites should be isolated and their biological characteristics (phenotype) defined using animal models and in vitro cell cultures. Ethnic background, travel history, dietary habits, medication and the presence of cats in the household should be added to the patients' data.

In our study 96.05% of the patients had acute infection. For this reason it is important that patients with toxoplasmosis infection are diagnosed and referred for treatment for the possibility of severe toxoplasmosis, especially after chemo- and radiotherapy when their immune system is weakened.

6.7. Molecular detection of *T. gondii*

We used standard *T. gondii* PCR markers (B1, SAG1, SAG2, SAG3, GRA6, APICO, L358, PK1, SAG1-Su, BTUB, alt.SAG2, c22-8 and c29-2), histological and immunohistochemical staining to confirm the presence of the parasite. All 88 tested samples were confirmed to be positive for *T. gondii* with markers B1, SAG1, SAG2 3/, SAG2 5/ and SAG3, giving a prevalence of 100% (95% CI: 95.82%). The ability to detect *T. gondii* genomic DNA in clinical samples made it possible to directly type the parasite and compare the results with the *T. gondii* database (htttp://toxodb.org/toxo). Only two samples were fully genotyped on all 10 loci and were identified as new strains. From all successfully genotyped samples, 34 had a single infection identified using all loci and 42 were identified as representing mixed infection at one or more loci with all three genotypes present. Type II genotype was the most predominant, followed by Type I and Type III. We detected 11 unusual genotypes. The presence of structures consistent with infection by the parasite was confirmed in 67 out of 76 tested haematoxylin-eosin stained tissue samples. With haematoxylin and eosin staining we can detect tachyzoites in alveolar cell linings, filling macrophages or free on the alveolar space and surrounded by inflammatory cells usually mononuclear. The main problem of this method is the detection of the parasite, especially in the areas where many host cells are present as they have the same colour after staining. Immunohistochemistry was performed in 76 tissue sections using polyclonal antibodies produced in rabbits. All 76 sections were confirmed to be positive for *T. gondii* showing different stages of the parasite and a wide range of infected cell types. A surprisingly high number of patients (96.05%) showed evidence of an active infection. Given that the background levels of infection in the UK are considered to be low at 10% and the global infection rate is 30% (Pappas *et* al 2009), this represents an extremely high infection rate in these cancer patients.

6.7. Genotypes of *T. gondii* detected in immunocompetent and immunocompromised patients

Human toxoplasmosis is mainly caused by Type II lineage, Type I and atypical strains are found in severe toxoplasmic retinochoroiditis and acute disseminated toxoplasmosis in immunocompromised patients. Moreover patients with ocular toxoplasmosis and animals are often infected by Type III strains (Howe *et al.* 1995). Genotypes in human asymptomatic population are largely unknown. In mice, type II (based on SAG2 PCR-RFLP) produce a high number of cysts and are more prone to reactivate in experimentally immunocompromised mice. It is possible that this type behaves the same in the human population but 15% of observed cysts in human patients belonged to type I, suggesting that type I can create cysts in human tissue (Howe *et al.* 1997). A few cases suggesting that the human immune system is less adapted to atypical genotypes than to genotypes usually circulating in domestic cycles (Darde *et al.* 1998). To date, no parasite strain has been shown to have specific pulmonary tropism, but there are some suggestions that South American strains could be more virulent, leading to disseminated toxoplasmosis. In our study Type II strains have been shown to be most prevalent and this corresponds to some previous studies on *T. gondii* in congenital infection and AIDS patients in North America and Europe (Ajzenberg *et al.* 2009). There is not enough data on *T. gondii* strains circulating in human populations in the UK to support or refuse our findings. Aspinall and her team (2002) analyzed 32 cases of human toxoplasmosis by PCR targeting the SAG2 gene. These consisted of toxoplasmic encephalitis in HIV/AIDS, 19 cases of congenital toxoplasmosis, three immunosuppressed patients and two immunocompetent patients. The majority of the clinical samples produced clear restriction patterns classifying them as type I, II and III, however some samples produced complex digestion patterns suggesting the samples were mix of SAG2 type I and II, and type I and III. The high frequency of type II strains may reflect the source of strains as latent infections of domestic and wild animals were often caused by types II and III (Howe *et al.* 1995). Type I strains have been associated with high level virulence in mice as they have the ability to cross epithelial barriers rapidly (Sibley *et al.* 1992). The unusually severe encephalitis and chorioretinitis associated with the reactivation of *T. gondii* cyst of a genotype I/III strain was described in an immunocompromised patient (Genot *et al.* 2007). In some other studies in France type II isolates showed a large predominance (>80%), with the occurrence of some type I and III isolates (Ajzenberg *et al.* 2002). In Poland, all genotyped samples were type II (Nowakowska *et al.* 2006). Also samples from Serbia and Romania among cerebral toxoplasmosis were type II (Djurkovic-Djakovic *et al.* 2006). In Europe, type III is more frequent in Mediterranean countries (Messaritakis *et al.* 2008) and it has been reported in Portugal (Waap *et al.* 2008), Spain (Fuentes *et al.* 2001) and Serbia (Djurkovic-Djakovic *et al.* 2014). Concurrent infections with parasites of different *T.gondii* types have been reported only in a few cases. Our finding that 47 out of 88 infections were mixed was unexpected. Only a small number (less than 5%) of isolates have mixtures of the two-allele patterns seen in the type strains, indicating that they are natural recombinants. It has recently been suggested that a wider range of genotypes might be found in geographically remote areas where transmission may rely more often on the sexual phase (Ajzenberg *et al.* 2004) which only occurs in cats. It is hard to explain such a high frequency of mixed infections in our study as a consequence of exposure to mixed oocysts. One possible explanation is via food born infection on multiple occasions of more than one *T. gondii* type during their life. In previous work Aspinall *et al.* (2002) tested commercial meat products and 38% of samples were *T. gondii* positive containing both SAG2 Type I and II mixtures. This scenario could account for the high number of mixed infections identified in our study. *Toxoplasma gondii* infection was considered to be acquired in the early stage and the prevalence is enhanced with age and declines in later stage (Nimir *et al.* 2010) and the average age of our patients was 69.65. Interestingly in a study from China (Zhang *et al.* 2015) a higher prevalence was in the younger population compared to the older age group. There was no significant difference in prevalence between males and females. We do not have the data on ethnic background of the patients or their travel history. It is possible that they travelled abroad or moved from a different country and brought back rare strains with them. Selected samples for each marker were sent for sequencing to confirm the mixed infections and resolve the strain type. Samples showing mixture of two or three alleles revealed possible double peaks at known polymorphic sites indicating the presence of more alleles in the same sample. Due to the presence of strong background noise we cannot definitely confirm or exclude the presence of mixed strains. Sequence analysis has as well limited capacity to detect mixtures where one component is in significant excess over another.

6.8. Molecular epidemiological studies of *T. gondii*

Our study has shown that the analysis with only one or three markers is not sufficient and could easily misidentify the strains as representing a simple set of genotypes or single genotype. However most of the highly repetitive and sensitive markers (e.g. B1, SAG1) are not useful for genotyping the parasite due to a relative absence of polymorphism. Typing of a single allele (one locus) does not necessarily imply that all alleles are of the same type. The three clonal lineages *T. gondii* system is based on the results of SAG2 typing and only recently after new multilocus typing was used more unusual strains started to emerge. Mixed infections can be more common in some regions and these strains may present different clinical complications. In our study we used direct genotyping of clinical samples while all of the previously analyzed strains were isolated by repeated passage in mice and in vitro culture, and therefore the apparent frequencies of specific lineages may have been biased against strains that were more difficult to isolate. The study of Costa *et al.* (1997), the first to be performed directly on biological samples (37 amniotic fluid samples) reported no mixed infections by analysis of microsatellites in the gene encoding *T. gondii* β-tubulin. In another study (Vilenna *et al.* 2004), samples from tested amniotic fluids, only one had mixed infection (Type I and II) detected by direct typing but after the murine inoculation, only Type II strain was isolated. Direct microsatellite typing on brain tissue from infected pigeons without previous passage in mice appeared to be better and in agreement with genotypes determined by bioassay (Waap *et al.* 2008). The presence of mixed strains infections in primary samples could demonstrate a possible selection bias in following parasite inoculation into mice. Due to the limited sample amount and technical difficulties of direct genotyping we were not able successfully amplify and type all loci. A high number of samples were positive at the beginning of the project but were PCR-negative with other markers at the end of the study. Repeated freezing and thawing could lead to a DNA damage and deterioration of sample quality. It has been suggested to add 50% glycerol to the samples to prevent DNA shearing due to ice crustal formation (Roder *et al.* 2010). With the development of new molecular methods it will be possible to study correlation between the parasite genotype and disease patterns in infected patients. More detailed strain comparisons based on wider set of sequence based markers will be necessary the global population structure of *T. gondii*. There is an interest in determining how particular genotypes might differ in their capacity to induce pathology or occurrence in particular host species. Not all seropositive AIDS patients develop encephalitis and similarly seroconversion during pregnancy does not always lead to infection of the foetus. Different migration capacities might contribute to strain specific differences in parasite dissemination. Dellacasa-Lindberg *et al.* (2007) demonstrated that a subpopulation of Type I parasite display enhanced migration in vitro compared with types II and III. One tachyzoite of Type I strain is sufficient to generate high parasite loads. Most *T. gondii* strains have been genotyped with only a small number of markers that are shared among many different strains. Understanding the genetic factors that influence the virulence and the mechanisms of genotype selection according to host species could contribute to the development of therapeutics designed to eliminate transmission or cure disease. The diagnosis of *T. gondii* infection is usually based on histopathological examination, serological assay and isolation of the parasite by mouse inoculation. Molecular assay such as PCR make it possible to detect small quantities of DNA and provide a sensitive diagnostic tool. In our study we performed different methods, molecular and histopathological for detecting *T. gondii*. We were not able obtain blood from mice and humans and conduct serological studies on our samples. No immunohistochemistry was performed on the mice samples so the nested PCR was the only method of detection. The main purpose of this study was to design a specific and sensitive nested PCR assay to detect the parasite in human lung tissue sample and brain tissue from wood mice. The nested PCR followed by RFLP assay used in our study is a fast and highly sensitive method for detecting *T. gondii* and characterizing parasite directly on clinical samples. Although the isolation of the parasite in mice is considered the gold standard, the technique is time consuming and we worked on some old samples, stored in lysis buffer or at -20˚C. The demonstration of the parasite after the inoculation is often delayed since mice usually develop a latent rather than an acute infection (Derouin *et al.* 1987). It could take one or two months of repeated passages to obtain a sufficient amount of tachyzoites and this could be difficult to achieve especially in slow dividing strains (Darde *et al.* 1996). It would be interesting compare the seroprevalence for both tested species and as well as histopathological findings in brains and lung of mice and humans. In the work carried out by Hermann *et al.* (2012) the serological detection resulted in higher prevalence when compared with molecular methods. Our prevalences in wood mice correspond with some previously reported data, but our human lung tissue samples have the highest *T. gondii* prevalence ever reported. Lower values for the PCR detection in some study samples are possibly due to the irregular distribution of the parasite in the tissue. This was confirmed with immunostaining when different stages of the parasite and different intensities were observed. The DNA from whole mice brain was used for detection of *T. gondii* and the large amount of host DNA could inhibit the parasite specific PCR amplification. In contrast, from some lung samples only a small amount of tissue was used for the DNA extraction, as the remaining part was used for histology. Maybe testing a greater number of tissue samples from each species by PCR could increase the overall detection rates. In some cases where multiple tissues were collected from individual animals, that more than one tissue could was positive for *T. gondii*. For example in study conducted by Burrells *et* al (2013), selections of badger tissues were positive for *T. gondii*, while the corresponding brain tissue was negative.

6.9. Markers for individual characterization of *T. gondii*

Another limitation of the markers used in our study is that they are not able to distinguish closely related genotypes derived from the same clonal lineage. Microsatellite typing is now considered one of the most powerful tools to study DNA polymorphism and is used often with combination with RFLP markers. A total of 15 microsatellite markers (TUB2, W35, TgM-A, B17, B18, M33, IV.1, XI.1, M48, M102, N60, N82, AA, N61 and N83) have been used to genotype *T. gondii.* Microsatellites are especially useful revealing recent mutations in closely related isolates within the lineage and can detect mixed infections. Since *T. gondii* is haploid, only one peak is expected for a given locus corresponding to one allele. More than one peak will be detected if mixed infections with different alleles are present in the sample (Ajzenberg *et al.* 2002). Microsatellite PCR typing requires only a small amount of DNA and allele sizing can be achieved with automatic sequencing by using fluorescent primers. The usefulness of microsatellite markers for isolate typing was demonstrated by Costa *et al.* (1997), Blacskton *et al.* (2001) and Ajzenberg *et al.* (2002). The limitation of this assay is the requirement of an automated sequencer and the presence of the host DNA can prevent detection of specific peaks or generate low peak intensity. As not sufficient amount of purified *T. gondii* DNA was obtained, this method was not performed on tested samples.

Traditional detection of *T. gondii* from environmental and tissue samples relied on microscopic examination, which is often of low sensitivity and thus unreliable. Different stages of the parasite can be stained with Haematoxylin and Eosin or Giemsa stain, which is cheap and time-effective. Its problem is the detection of tachyzoites which can be easily missed especially if many host cells are present. In one study of cerebral toxoplasmosis the organism was not seen in over 50% of positive cases (Luft *et al.* 1984). Periodic acid Schiff (PAS) is used to stain granules in bradyzoites (Dubey *et al.* 1998). This method is time consuming and the parasite can be easily missed if only the tachyzoite stage is present, especially in the early stages of the infection. Electron microscopy is rarely used due its difficulty for routine use.

6.10 Future work and recommendations

The research presented in this thesis has addressed many questions regarding the prevalence and genotypes of *T. gondii* in natural populations of rodents and human, but there are still several lines that should be pursued.

6.10.1. *Toxoplasma gondii* in wood mouse

As discussed in Chapter 3, the selection of ten *A. sylvaticus* microsatellites was sufficient enough to successfully analyse the population structure of the wood mice. However two of the selected microsatellites showed very small variation and so replacing them with more polymorphic ones could achieve better results and more accurate analysis of the rodent population. Some improvements could be also done regarding the sample size. Our original sample size had 126 wood mice, which was enough to identify four populations but it was insufficient to analyse the mice on family levels. In our sampling site we have identified 22 families with various numbers of siblings, ranging from two up to ten members in each family. To increase the statistical power and to identify more siblings from individual family the minimum sample size needs to be increased. Using modelling this estimated that a sample size with minimum of 500 animals is required, if we need 22 families with 20 members in each family. This requires more regular sampling over the sampling period. Ideally the sampling should be over a one year period, starting in January and finishing in December, as this will ensure better sample set for use in COLONY as mice are more likely to be same generation siblings rather than transgeneration. The mice should be collected preferably twice a month, from the same sampling sites. If this is done the population data from this study can be used to estimate the number of rodents in the individual sampling areas. Based on the allele frequencies (0.05), NeEstimator software estimated that 134.6 wood mice from the R1 population, 136.2 from the R2, 55 from the G and 227.2 from the B population are present in this area. The software was not able to estimate the number of mice for the whole area. Possible replacement of two microsatellites could solve this problem and give us a more reliable overall estimate. Only 24 of the collected mice belonged to the population B. This is a much smaller sample size, when compared for example population G, where 40 mice were assigned. To make the identification of the family units more accurate, it is necessary to have more balanced sample sizes. In Malham Tarn Nature Reserve there are a few more sites, which could possibly have large mouse populations, for example local farms, other areas in the forest or by expanding further south and sampling mice around Malham Cove. As wood mice are quite prevalent in the UK and inhabit variety of habitats it would be interesting to include new sampling areas outside the Malham Tarn Nature Reserve. The aim of this expansion would be to be able to extend the outcomes of this work from a local to a national or international setting.

In Chapter 4 the prevalence and genotypes of *T. gondii* in natural populations of wood mice were analysed in relation to their population structure. The results showed an unusually high prevalence of the parasite in an area relatively free of cats. Surprisingly a high number of mice had mixed infections, harbouring two or all three main genotypes. This indicates that mixed infections are more frequent than previously thought. We do not know what the original source of the infection is and so can only speculate. The mice probably became infected by consuming oocyst together with infected food (e.g. carrion) or vertically transmitted from mother to offspring. Most likely it is the combination of all of these routes. *T. gondii* was detected in 44 out of 126 tested mice belonging to four different populations (Chapter 4). There was a significant difference between each genetic population. The direct genotyping method was used and it was possible to successfully genotype all positive samples on all tested loci (SAG2, SAG3 and GRA6). As not all standard *T. gondii* markers worked on these samples it was not possible to match our detected strains with those listed in the database (www. toxodb.org). In future studies it is recommended the full set of markers be used. One of the major issues was the presence of mixed infections in high number of mice. The identification of the strains from the RFLP gel images was very difficult to interpret due to the presence of partial digests or complex binding patterns and DNA smears. Using direct PCR RFLP it was not possible to obtain a sufficient volume of purified PCR to be sent for sequencing. One of the possible solutions for the strain identification is the isolation of parasites directly from the tissue. The direct genotyping of cultured isolates can be also difficult if two or more strains are present in the culture. This is especially true if there is the different growth rate of individual strains in the medium. For example the RH (Type I) strain divides rapidly and can overgrown other strains in a very short time, what can lead to their disappearance from the medium and so genotyping results may not reflect the real strain structure of the isolate. It is also very difficult to grow some strains in cell culture. A possible solution might be a greater variety media to encourage better growth conditions for different strains. However, as yet there is no background data on this issue and with a current number of 231 *T. gondii* genotypes; this would be time consuming task. However better analysis of the isolates could help improve the way we address the challenging questions regarding the parasite transmission. An attempt was made to find any relationship between the infection and family structure to determine whether the same genotypes are present in the same families and populations. Due to the small number of siblings in each family and the presence of mixed infections, it was not possible to fully answer this question although we have found that some families shared the same genotypes. A larger sample size and better identification of the genotypes is necessary to improve this analysis. Another option on how to perform genotyping with a limited amount of sample or the parasite is uncultivable is to develop whole genome amplification to enrich for *T. gondii* DNA where the amplified DNA can be directly used in a range of genetic analyses (qPCR, sequencing, microsatellite analyses, etc.). One of the disadvantages of the whole genome typing approach is the risk of multiplying host or bacterial DNA and obtaining poor resolution if mixed genotypes are present. Based on all of the above mentioned techniques the isolation and cultivation of the parasite seems to be a crucial step. Before the parasite isolation it would be favourable to perform fast screening tests for the presence of *T. gondii* (e.g. IHC, Haematoxylin and eosin staining, LAMP, serological screening) as it would be time consuming and expensive to try to extract the parasite from few hundred samples from which would be possibly negative. After successful isolation and cultivation of different *T. gondii* strains it would be useful to set up a biobank where researchers from different areas could retrieve the data for their studies. Similar collections could be set up for parasites obtained from different tissue types and animals, including humans.

6.10.2. *T. gondii* in lung cancer patients

Data presented in Chapter 5 show that *T. gondii* was present in cells of 88 lung cancer tissue using molecular, histological and immunohistological methods. This 100% prevalence is in contrast to the average 30% worldwide infection rate. *T. gondii* markers B1, SAG1, SAG2 and SAG3 provided fast and reliable detection method and additionally can be used for strain typing. Future work would be aimed at gaining a full strain identification using a full set of markers as described in previous chapters. Immunohistochemistry and haematoxylin and eosin staining was used as a secondary confirmation method. During this research two main issues emerged. A similar problem was observed as found as with the wood mice samples using the direct genotyping method. From 88 tested samples, only two could be fully genotyped on all tested loci. We were not able to obtain enough of purified PCR product from the sample heavily biased towards host DNA. Again, the solution for this problem could be the isolation and cultivation of the parasite. Genotyping performed on the isolates could solve the problem in identifying mixed infections present in tested patients. The second main issue was the lack of negative control samples from subjects without lung cancer. Although few negative controls were used for the PCRs and IHC (negative mouse tissue, cell culture) it was not possible able to test our methods on lung tissue from healthy individuals. Our modelling based on 100% infection rate showed that we need a minimum of 20 samples from healthy individuals (no cancer, COPD or other underlying conditions) to make it statistically robust. These control samples should have balanced by sex and age and preferably tested subjects should be free of any medication. Unfortunately it is very difficult to obtain lung tissue from healthy donors. In an ideal world a few different control groups could be obtained to control possible link between *T. gondii* and cancer or other diseases. In some studies they successfully decreased the size of the tumours using *T. gondii* lysate (Chapter 5) while in others they describe the parasite as the source of cancer. It is possible that different strains of *T. gondii* may evoke different responses depending on the health status of the human host while many others, presently unknown, factors could play crucial role. To investigate this problem it would be necessary to take a different and larger scale approach. In the first place is important to collect more data about the prevalence and strains circulating in the human population, from both healthy individuals and from patients with a range of diseases, including different types of cancer. Subjects should be tested using serological, molecular, histological and immunohistological methods in addition to live parasitic culture. For epidemiological studies is also important to know the amount of the parasite present in tested tissues, which could be measured using qPCR. Samples need to be collected from different geographical locations across the country, from both, urban and rural area to obtain a full picture. A possible future development would be to set up biobanks with the collections of possibly all known strains and investigate their properties and virulence in different hosts, starting with the cell cultures and moving to laboratory animals later. Attention needs to be paid to the development of more diagnostic methods and markers; especially those to able detect virulent strains much faster. In our study we tried to use Scanning electron microscopy (SEM) and Transmission electron microscopy (TEM) to detect the presence of the parasite. These methods are expensive; time consuming and not very practical for routine use but still can be used for final confirmation. If these methods are used on the tissue samples where not many host cells are present, they can provide another useful detection tool. They can be also used in the treatment analysis as they can show the effect of tested drugs on ultrastructural level.

Different types of tissue may have different susceptibility to parasites and this could explain the results in our study. Finally there were differences in status of tested subjects, where the human patients were already immunocompromised and so at risk contracting toxoplasmosis while the wild mice were at relatively good health as far as we are aware. What the two tested groups had in common was an unusually high number of mixed infections and the dominance of Type I. These results raise some new questions. What is the actual impact of individual strains on our tested subjects? We can estimate the effect of *T. gondii* on the cancer patients as there are some studies available, but there is little data on how the parasite affects wild animals. Only a few reports mention the negative impact of *T. gondii* on sea otters and dolphins, but most of the data is related to farm animals (e.g. pigs, sheep and chickens) and the possible risk to humans. As mentioned previously, we need to learn more about parasite genetics, host genetics and mechanisms of virulence. This will be only possible by detailed analysis of as many possible strains from many different intermediate hosts including humans from a variety of habits from all continents. Recently many parasites and bacteria have started showing increased resistance to treatment and medication. At the moment there are no reports of resistant *T. gondii* strains but we cannot exclude this possibility. This could have fatal consequences especially for the patients with impaired immune systems or if the disease is caused by some highly virulent strain. Many of the European countries run screening programmes for pregnant women but no automatic tests are available in the UK (until requested) and many of the pregnant women are not aware of the danger that *T. gondii* could cause to their unborn child. More education needs to be provided to larger communities and GP surgeries as *T. gondii* is often overlooked and ignored. In longer term, the founding of international networks of reference laboratories and information exchange between experts, doctors and ecologists could bring more answers to these outstanding questions.

References

1. Albdul-Muneer, P.M. (2014). Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. Genetics Research International. 2014: 691759.

2. Ajzenberg, D.; Cogné, N.; Paris, L.; Bessières, M.H.; Thulliez, P.; Filisetti, D.; Pelloux, H.; Marty P.; Dardé, M. L. (2002). Genotype of 86 *Toxoplasma gondii* isolates associated with human congenital toxoplasmosis, and correlation with clinical findings. Journal of Infectious Diseases. 186(5): 684 – 689.

3. Afonso, E.; Poulle, M.L.; Lemoine, M.; Villena, I.; Aubert, D. and Gilot-Fromont, E. (2007). Prevalence of *Toxoplasma gondii* in small mammals from the Ardennes region, France. Folia Parasitologica (Praha). 54 (4): 313 – 314.

4. Ahmadpour, E.; Daryani, A.; Sharif, M. (2014). Toxoplasmosis in immunocompromised patients in Iran: a systematic review and meta-analysis. Journal of Infection in Developing Countries. 8(12):1503-1510.

5. Araujo, F.G.; Slifer, T. (2003). Different strains of *Toxoplasma gondii* induce different cytokine responses in CBA/Ca mice. Infection and Immunology. 71(7): 4171-4174.

6. Aspinall, T.; V.; Marlee, D.; Hyde, J.E.; Sims, P.F. (2002). Prevalence of *Toxoplasma gondii* in commercial meat products as monitored by polymerase chain reaction--food for thought? International Journal of Parasitology. 32(9): 1193-1199.

7. Aspinall, T.; V.; Guy, E.C.; Roberts, K.E, Joynson, D.H.; Hyde, J.E.; Sims, P.F. (2003). Molecular evidence for multiple *Toxoplasma gondii* infections in individual patients in England and Wales: public health implications. International Journal of Parasitology. 33(1): 97-103.

8. Attuquayefio, D.K.; Gorman, M.L.; Wolton, R. J. (1986), Home range sizes in the Wood mouse *Apodemus sylvaticus*: habitat, sex and seasonal differences. Journal of Zoology. 210 (1): 45–53.

9. Dellacasa-Lindberg, I.; Hitziger. N.; Barragan, A. (2007). Localized recrudescence of *Toxoplasma* infections in the central nervous system of immunocompromised mice assessed by in vivo bioluminescence imaging. Microbes and Infection. 9(11): 1291-1298.

10. Barta, J.R. (2001). Molecular approaches for inferring evolutionary relationships among protistan parasites. Veterinary Parasitology. 101 (3-4): 175 – 186.

11. Barry, M.A. (2013). Childhood parasitic infections endemic to the United States. Paediatric Clinics of North America. 60 (2): 471-485.

12. Baum, J.; Tonkin, C.J.; Paul, A.S.; Rug, M.; Smith, B.J.; Gould, S.B.; Richard, D.; Pollard, T.D.; Cowman, A.F. (2008). A malaria parasite forming regulates actin polymerization and localizes to the parasite-erythrocyte moving junction during invasion. Cell Host Microbe. 3:188–198.

13. Beasley, V.; Joshi, P.V.; Singanayagam, A.; Molyneaux, PL.; Johnston, SL.; Mallia, P. (2012). Lung microbiology and exacerbations in COPD. International Journal of Chronic Obstructive Pulmonary Disease. 7:555-569.

14. Benenson, M.W.; Takafuji, E.T.; Lemon, S.M.; Greenup, R.L.; Sulzer, A.J. (1982). Oocyst-transmitted toxoplasmosis associated with ingestion of contaminated water. New England Journal of Medicine. 307: 666–669.

15. Bengtson, S.A.; Nilsson, A.; Rundgren, S. (1989). Population structure and dynamics of wood mouse *Apodemus sylvaticus* in Iceland. Holarctic Ecology. 12: 351-368.

16. Berger, J. (1966). Serologic research on *Toxoplasma* infections in animals in the area of the State Veterinary Diagnostic Laboratory Frankfurt/Main. Deutsche Tierarztliche Wochenschrift 73 (11): 261-263.

17. Berger-Schoch, A.E.; Bernet, D.; Doherr, M.G.; Gottstein, B.; Frey, C.F. (2011). *Toxoplasma gondii* in Switzerland: a serosurvey based on meat juice analysis of slaughtered pigs, wild boar, sheep and cattle. Zoonoses Public Health. 58 (7): 472-478.

18. Berger-Wolf, T.Y.; Sheik, S.I.; DasGupta, B.; Ashley, V.M.; Caballero, I.c.; Putrevu, S.L. (2007). Reconstructing sibling relationships in wild populations. Bioinformatics. 23(13): 49-56.

19. Beverley, J.K.A. (1959). Congenital transmission of toxoplasmosis through successive generation of mice. Nature 183 (4671): 1348-1349.

20. Black, M.W.; Boothroyd J.C (2000). Lytic cycle of *Toxoplasma gondii*. Microbiology and Molecular Biology Reviews. 64 (3): 607-623.

21. Blackston, C.R.; Dubey, J.P.; Dotson, E.; Su, C.; Thulliez, P.; Sibley, D.; Lehmann, T. (2001). High-resolution typing of *Toxoplasma gondii* using microsatellite loci. Journal of Parasitology. 87 (6): 1472-1475.

22. Bowie, W.R.; King, A.S.; Werker, D.H.; Isaac-Renton, J.L.; Bell, A.; Eng, S.B.; Marion, S.A. (1997). Outbreak of toxoplasmosis associated with municipal drinking water. The BC Toxoplasma Investigation Team. Lancet. 350 (9072): 173-177.

23. Boyce, K.; Hide G.; Craig, P.S.; Harris, P. D.; Reynolds, C.; Pickles, A. and Rogan, M.T. (2012). Identification of a new species of digenean *Notocotylus malhamensis* n. sp. (Digenea: Notocotylidae) from the bank vole (*Myodes glareolus*) and the field vole (*Microtus agrestis*). Parasitology. 139 (12): 1630-1639.

24. Boyce, K.; Hide, G.; Craig, P.S.; Reynolds, C.; Hussain, M.; Bodell, A.J.; Bradshaw, H.; Pickles, A. and Rogan, M.T. (2013). A molecular and ecological analysis of the trematode *Plagiorchis elegans* in the wood mouse *Apodemus sylvaticus* from a periaquatic ecosystem in the UK. Journal of Helminthology. 88 (3): 1-11.

25. Brinkmann, B.; Klintschar, M.; Neuhuber, F.; Huhne, J.; Rolf, B. (1998). Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. American Journal of Human Genetics. 62(6): 1408-1415.

26. Burrells, A.; Bartley, P.M.; Zimmer, I.A.; Roy, S.; Kitchener, A.C.; Meredith, A.; Wright, S.E.; Innes, E.A.; Katzer, F. (2013). Evidence of the three main clonal *Toxoplasma gondii* lineages from wild mammalian carnivores in the UK. Parasitology. 140(14): 1768-1776

27. Butet, A. (1994). Nutritional conditions and annual fluctuations in *Apodemus sylvaticus* populations. Russian Journal of Ecology. 25 (3): 111-119.

28. Callen, D.F.; Thompson, A.D.; Shen, Y.; Phillips, H.A.; Richards, R.I.; Mulley, J.C.; Sutherland, G.R. (1993). Incidence and origins of “null” alleles in (AC)n microsatellite markers. American Journal of Human Genetics. 52 (5): 922-927.

29. Carme, B.; Bissuel, F.; Ajzenberg, D.; Bouyne, R.; Aznar, C.; Demar, M.; Bichat, S.; Louvel, D.; Bourbigot, A.M.; Peneau, C.; Neron, P.; Dardé, M.L. (2002). Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. Journal of Clinical Microbiology. 40(11): 4037-4044.

30. Catterall, J.R.; Hofflin, J.M.; Remington, J.S. (1986). Pulmonary toxoplasmosis. American Review of Respiratory Disease. 133(4): 704–705.

31. Chiappino, M.L, Nichols, B.A, O’Connor, G.R. (1984). Scanning electron microscopy of *Toxoplasma gondii:* parasite torsion and host-cell responses during invasion. Journal of Protozoology. 31: 288–292.

32. Chinchilla, M, Ruiz, A. (1976). Cockroaches as possible transport hosts of *Toxoplasma gondii* in Costa Rica. Journal of Parasitology. 62(1): 140-142.

33. Choo, J.D.; Lee, J.S.; Kang, J.S.; Lee, H.S.; Yeom, J.Y.; Lee, Y.H. (2005). Inhibitory Effects of *Toxoplasma* Antigen on Proliferation and Invasion of Human Glioma Cells. Journal of Korean Neurosurgical Society. 37(2): 129-136.

34. Cohen, W.; Koslow, M. (1985). An unusual CT presentation of cerebral toxoplasmosis. Journal of Computer Assisted Tomography. 9(2): 384–386.

35. Cornuet, J.M.; Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics. 144(4): 2001-2014.

36. Cornuet, J.M.; Pudlo, P.; Veyssier, J.; Dehne-Garcia, A.; Gautier, M.; Leblois, R.; Marin, J.M.; Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 15(30): 1187-1189.

37. Costa, J.M.; Dardé, M.L.; Assouline, B.; Vidaud, M.; Bretagne, S. (1997). Microsatellite in the beta-tubulin gene of *Toxoplasma gondii* as a new genetic marker for use in direct screening of amniotic fluids. Journal of Clinical Microbiology. 35(10): 2542-2545.

38. da Silva, A.V.; Langoni, H. (2001). The detection of *Toxoplasma gondii* by comparing cytology, histopathology, bioassay in mice, and the polymerase chain reaction (PCR) Veterinary Parasitology. 97(3): 191–198.

39. Dabritz, H.A.; Miller, M.A.; Gardner, I.A.; Packham, A.E.; Atwill, E.R.; Conrad, P.A. (2008). Risk factors for *Toxoplasma gondii* infection in wild rodents from central coastal California and a review of *T. gondii* prevalence in rodents. Journal of Parasitology 94(3): 675-683.

40. Darani, H.Y.; Shirzad, H.; Mansoori, F.; Zabardast , N.; Mahmoodzadeh, M. (2009). Effects of *Toxoplasma gondii* and *Toxocara canis* antigens on WEHI-164 fibrosarcoma growth in a mouse model. Korean Journal of Parasitology. 47(2): 175-177.

41. Darani, H.Y.; Yousefi, M. (2012). Parasites and cancers: parasite antigens as possible targets for cancer immunotherapy. Future Oncology. 8(12): 1529-35.

42. Dardé, M.L.; B. Bouteille, M. Pestre-Alexandre. (1988). Isoenzymic characterization of seven isolates of *Toxoplasma gondii* by isoelectrofocusing in polyacrylamide gels. American Journal of Tropical Medicine and Hygiene. 39(6): 551-558.

43. Dardé, M.L.; Villena, I.; Pinon, JM.; Beguinot, I. (1998). Severe toxoplasmosis caused by a *Toxoplasma gondii* strain with a new isoenzyme type acquired in French Guyana. Journal of Clinical Microbiology. 36(1): 324.

44. Daval, S.; Poirier, P.; Armenaud, J.; Cambon, M.; Livrelli, V. (2010). Development of a real-time PCR assay for quantitative diagnosis of *Toxoplasma gondii* after allogeneic bone marrow transplantation. Pathologie Biologie (Paris). 58(1): 104-109.

45. de Marco, R.; Accordini, S.; Cerveri, I.; Corsico, A.; Antó, J.M.; Künzli, N.; Janson, C.; Sunyer, J.; Jarvis, D.; Chinn, S.; Vermeire, P.; Svanes, C.; Ackermann-Liebrich, U.; Gislason,. T.; Heinrich, J.; Leynaert, B.; Neukirch, F.; Schouten, J.P, Wjst, M.; Burney, P. (2007). Incidence of chronic obstructive pulmonary disease in a cohort of young adults according to the presence of chronic cough and phlegm. American Journal of Respiratory and Critical Care Medicine. 175(1): 32-39.

46. Derouin, F.; Mazeron, M.C.; Garin, Y.J. (1987). Comparative study of tissue culture and mouse inoculation methods for demonstration of *Toxoplasma gondii*. Journal of Clinical Microbiology. 25(9): 1597-600.

47. de Sousa, S.; Ajzenberg, D.; Canada, N.; Freire, L.; da Costa, J.M.; Dardé, M.L.; Thulliez, P.; Dubey, J.P. (2006). Biologic and molecular characterization of *Toxoplasma gondii* isolates from pigs from Portug*al*. Veterinary Parasitology. 135(2): 133-136.

48. Di Stefano, A.; Turato, G.; Maestrelli, P.; Mapp, C.E.; Ruggieri, M.P.; Roggeri, A.; Boschetto, P.; Fabbri, L.M.; Saetta, M. (1996). Airflow limitation in chronic bronchitis is associated with T-lymphocyte and macrophage infiltration of the bronchial mucosa. American Journal of Respiratory and Critical Care Medicine. 153(2): 629-632.

49. Dib, C. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380(6570): 152-154.

50. Dickman, C.; Doncaster, C. (1987). The ecology of small mammals in urban habitats. Populations in a patchy environment. Journal of Animal Ecology, 56(2): 629-640.

51. Dietrich, W. (1996). A comprehensive genetic map of the mouse genome. Nature 380(6570): 149-152.

52. Djurković-Djaković, O.; Klun, I.; Khan, A.; Nikolić, A.; Knezević-Usaj, S.; Bobić, B.; Sibley, L.D. (2006). A human origin type II strain of *Toxoplasma gondii* causing severe encephalitis in mice. Microbes and Infection. 8(8): 2206-2212.

53. Do, C.; Waples, R. S.; Peel, D.; Macbeth, G. M.; Tillett, B. J.; Ovenden, J. R. (2014). NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Molecular Ecology Resources. 14(1): 209-214.

54. Dodd, N.S.; Lord, J.S.; Jehle, R.; Parker, S.; Parker, F.; Brooks, D.R. and Hide, G. (2014). *Toxoplasma gondii*: Prevalence in species and genotypes of British bats (*Pipistrellus pipistrellus* and *P. pygmaeus*). Experimental Parasitology 139: 6-11.

55. Dubey, J.P.; Miller, N.L.; Frenkel, J.K. (1970a). The *Toxoplasma gondii* oocyst from cat faeces. Journal of Experimental Medicine. 132(4): 636–662

56. Dubey, J.P.; Frenkel, J.K.; (1972). Cyst-induced toxoplasmosis in cats. Journal of Protozoology. 19(1): 155–177.

57. Dubey, J.P.; Frenkel, J.K. (1976). Feline toxoplasmosis from acutely infected mice and the development of *Toxoplasma* cysts. Journal of Protozoology. 23(4): 537–546.

58. Dubey, J.P. (1977). *Toxoplasma*, *Hammondia*, *Besnoitia*, *Sarcocystis* and other tissue cyst-forming coccidia of man and animals. In: Kreier, J.P. (Ed.), Parasitic Protozoa vol. 3. Academic Press. pp. 101–237.

59. Dubey, J.P.; Beattie, C.P. (1988). Toxoplasmosis of Animals and Man. CRC Press, Boca Raton, Fl. pp. 1–220.

60. Dubey, J.P.; Weigel, R.M.; Siegel, A.M.; Thulliez, P.; Kitron, U.D.; Mitchell, M.A.; Mannelli, A.; Mateuspinillia, N.E.; Shen, S.K.; Kwok, O.C.H. and Todd, K.S. (1995). Sources and reservoirs of *Toxoplasma gondii* infection on 47 swine farms in Illinois. Journal of Parasitology 81(5): 723-729.

61. Dubey, J.P.; Shen, S.K.; Kwok, O.C. and Thulliez, P. (1997). Toxoplasmosis in rats (*Rattus norvegicus*): congenital transmission to first and second generation offspring and isolation *Toxoplasma gondii* from seronegative rats. Parasitology 115(1): 9-14.

62. Dubey, J.P.; Lindsay, D.S.; Speer, C.A.; (1998). Structure of *Toxoplasma gondii* tachyzoites, bradyzoites and sporozoites, and biology and development of tissue cysts. Clinical Microbiology Reviews. 11(2): 267–299

63. Dubey, J.P. (2001). Oocyst shedding by cats fed isolated bradyzoites and comparison of infectivity of bradyzoites of the VEG strain *Toxoplasma gondii* to cats and mice. Journal of Parasitology. 87(1): 215-219.

64. Dubey, J.P. (2005). Unexpected oocyst shedding by cats fed *Toxoplasma gondii* tachyzoites: In vivo stage conversion and strain variation. Veterinary Parasitology. 133(4): 289–298.

65. Dubey, J.P.; Jones, J.L. (2008). *Toxoplasma gondii* infection in humans and animals in the United States. International Journal of Parasitology 38(11): 1257-1278.

66. Duncanson, P.; Terry, R.S.; Smith, J.E. and Hide, G. (2001). High levels of congenital transmission of *Toxoplasma gondii* in a commercial sheep flock. International Journal for Parasitology 31(14): 1699-1703.

67. Eichenwald, H. (1948). Experimental toxoplasmosis. I. Transmission of the infection in utero and through the milk of lactating female mice. American journal of diseases of children. 76(3): 307 -315.

68. Elamin, M.H. (2014). Genotyping of *Toxoplasma gondii* from rats (*Rattus rattus*) in Riyadh, Saudi Arabia. Korean Journal of Parasitology. 52(3): 257-61.

69. Ellegren, H. (2000). Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genetics. 24: 400-402.

70. Endo, T.; Kobayashi, A. (1976). *Toxoplasma gondii* electron microscopic study on the Dye Test reaction. Experimental Parasitology. 40(2): 170-178.

71. Esteban-Redondo, I.; Maley, S.W.; Thomson, K.; Nicoll, S.; Wright, S.; Buxton, D.; Innes, E.A. (1999). Detection of *T. gondii* in tissues of sheep and cattle following oral infection. Veterinary Parasitology. 86(3): 155–171.

72. Elsaid, M.M.A.; Martins, M.S.; Frézard, F.; Braga, E.M. and Vitor RWA. (2001). Vertical toxoplasmosis in a murine model. Protection after immunization with antigens of *Toxoplasma gondii* incorporated into liposomes. Memoirs Institute Oswaldo Cruz 96(1): 99-104.

73. Evanno, G.; Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14(8): 2611–2620.

74. Evans, T.G.; Schwartzman, J.D. (1991). Pulmonary toxoplasmosis. Seminars in Respiratory Infections. 6(1): 51-57.

75. Fang, H.; Pengal, R.A.; Cao, X.; Ganesan, L.P.; Wewers, M.D.; Marsh, C.B.; Tridandapani, S. (2004). Lipopolysaccharide-induced macrophage inflammatory response is regulated by SHIP. Journal of Immunology. 173(1): 360-366.

76. Favoreto-Júnior, S.; Ferro, E.A.; Clemente, D.; Silva, D.A.; Mineo, J.R. (1998). Experimental infection of *Calomys callosus* (Rodentia, Cricetidae) by *Toxoplasma gondii.* Memórias do Instituto Oswaldo Cruz. 93(1): 103-107.

77. Flegr, J.; Klapilova, K.; Kankova, S. (2014). Toxoplasmosis can be a sexually transmitted infection with serious clinical consequences. Not all routes of infection are created equ*al.* Medical Hypotheses. 83: 289-289.

78. Flegr, J.; Havlícek, J.; Kodym, P.; Malý, M.; Smahel, Z. (2002). Increased risk of traffic accidents in subjects with latent toxoplasmosis: a retrospective case-control study. BMC Infectious Diseases. 2:11.

79. Fletcher. C.; Peto, R. (1977). The natural history of chronic airflow obstruction. British Medical Journ*al.* 1(6077): 1645-1648.

80. Flowerdew, J. R (1974). Field and Laboratory Experiments on the Social Behaviour and Population Dynamics of the wood mouse (*Apodemus sylvaticus*). Journal of Animal Ecology. 43(2): 499-511.

81. Frankham, R.; Ballou, J.D.; Briscoe, DA. (2002). Introduction to Conservation Genetics. Cambridge University Press. United Kingdom.

82. Franti, C.E.; Riemann, H.P.; Behmeyer, D.E.; Suther, D.; Howarth, J.A. and Ruppanner, R. (1976). Prevalence of *Toxoplasma gondii* antibodies in wild and domestic animals in northern California. Journal of the American Veterinary Medical Association. 169(9): 901-906.

83. Frenkel, J.K.; Dubey, J.P. and Miller, N.L. (1970). *Toxoplasma gondii* in cats: fecal stages identified as coccidian oocysts. Science. 167(3919): 893-896.

84. Frenkel, J.K. (1973). *Toxoplasma* in and around us. BioScience. 23(6): 343–352.

85. Freeman, K.; Tan, H.K.; Prusa, A.; Petersen, E.; Buffolano, W.; Malm, G.; Cortina-Borja, M.; Gilbert, R.; European Multicentre Study on Congenital Toxoplasmosis. (2008). Predictors of retinochoroiditis in children with congenital toxoplasmosis: European, prospective cohort study. Paediatrics. 121(5): 1215-1222.

86. Freyre, A.; Fialho, C.G.; Bigatti, L.E.; Araujo, F.A.; Falcón, J.D.; Mendez, J.; González, M. (2009). *Toxoplasma gondii*: congenital transmission in a hamster model. Experimental Parasitology. 122(2): 140-144.

87. Fuehrer, H. P.; Blosch, I.; Siehs, C. and Hass, A. (2010). Detection of *Toxoplasma gondii*, *Neospora caninum,* and *Encephalitozoon cuniculi* in the brains of common voles (*Microtus arvalis*) and water voles (*Arvicola terrestris*) by gene amplification techniques in western Austria (Vorarlberg). Parasitology Research. 107(2): 469 – 473.

88. Fuentes, I.; Rubio, J.M.; Ramírez, C. and Alvar, J. (2001). Genotypic Characterization of *Toxoplasma gondii* Strains Associated with Human Toxoplasmosis in Spain: Direct Analysis from Clinical Samples. Journal of Clinical Microbiology 39(4): 1566–1570.

89. Gagnon, P.; Guenette, J.A.; Langer, D.; Laviolette, L.; Mainguy, V.; Maltais, F.; Ribeiro, F.; Saey, D. (2014). Pathogenesis of hyperinflation in chronic obstructive pulmonary disease. International Journal of Chronic Obstructive Pulmonary Disease. 9: 187-201.

90. Gambashidze, K.; Khorava, P.; Azaladze, T.; Kalandarishvili, K.; Jaiani, E.; Lasareishvil, B.; Azaladze, A.; Tediashvili, M. (2012). Antitumor and adjuvant effects of phagelysates of *E.coli* in mice with Ehrlich carcinoma. Experimental Oncology. 34(2): 107-11.

91. Garcia, J.L.; Gennari, S.M.; Machado, R.Z.; Navarro, I.T. (2006). *Toxoplasma gondii*: detection by mouse bioassay, histopathology, and polymerase chain reaction in tissues from experimentally infected pigs. Experimental Parasitology. 113(4): 267–271.

92. Gebczynski M.; Ratkiewicz M. (1998). Does biotope diversity promote an increase of genetic variation in the bank vole population? Acta Theriologica. 43(2): 163-173.

93. Gelanew, T.; Kuhls, K.; Hurissa, Z.; Weldegebreal, T.; Hailu, W.; Kassahun, A.; Tamrat, A.; Abebe, T.; Hailu, A. and Schönian, G. (2010). Inference of population structure of *Leishmania donovani* strains isolated from different Ethiopian visceral leishmaniasis Endemic Areas. PLoS Neglected Tropical Diseases 4(11): e889.

94. Goldstein, D.B, Schlotterer, C. (1999). Microsatellites: Evolution and Applications. Oxford University Press, U.S.A.

95. Genot, S.; Franck, J.; Forel, J.M.; Rebaudet, S.; Ajzenberg, D.; de Paula, A.M.; Dardé, M.L.; Stein, A.; Ranque, S. (2007). Severe *Toxoplasma gondii* I/III recombinant-genotype encephalitis in a human immunodeficiency virus patient. Journal of Clinical Microbiology. 45(9): 3138-3140.

96. Gotteland, C.; Chaval, Y.; Villena, I.; Galan, M.; Geers, R.; Aubert, D.; Poulle, M.L.; Charbonnel, N. and Gilot-Fromont, E. (2014). Species or local environment, what determines the infection of rodents by *Toxoplasma gondii*? Parasitology 141(2): 259-268.

97. Green, R. (1979). The ecology of wood mice (*Apodemus sylvaticus*) on arable farmland. Journal of Zoology. 188(3): 357–377.

98. Grigg, M.E.; Ganatra, J.; Boothroyd, J.C. and Margolis, T.P. (2001). Unusual abundance of atypical strains associated with human ocular toxoplasmosis. Journal of Infectious Diseases. 184(5): 633-639.

99. Hafid, J.; Tran Manh Sung, R.; Raberin, H.; Akono, Z.Y.; Pozzetto, B.; Jana, M. (1995). Detection of circulating antigens of *Toxoplasma gondii* in human infection. American Journal of Tropical Medicine and Hygiene. 52(4): 336-339.

100. Hafid, J.; Flori, P.; Raberin, H.; Tran Manh Sung, R. (2001). Comparison of PCR, capture ELISA and immunoblotting for detection of *Toxoplasma gondii* in infected mice. Journal of Medical Microbiology. 50(12): 1100-1104.

101. Hall SM, Pandit A, Golwilkar A, Williams T.S. (1999). How do Jains get *Toxoplasma* infection? Lancet. 354(9177): 486-487.

102. Harr, B.; Musolf, K.; Gerlach, G. (2000). Characterization and isolation of DNA microsatellite primers in wood mice (*Apodemus sylvaticus*, Rodentia). Molecular Ecology. 9(10): 1664-1665.

103. Harris, S.; Morris, P.; Wray, S.; Yalden, D. (1995) A Review of British Mammals: Population Estimates and Conservation Status of British Mammals other than Cetaceans. Joint Nature Conservation Committee, Peterborough.

104. Hejlicek, K.; Literak, I. and Nezval, J. (1997). Toxoplasmosis in wild mammals from the Czech Republic. Journal of Wildlife Diseases 33(3): 480-485.

105. Hejlicek, K. and Literak, I. (1998). Long-term study of *Toxoplasma gondii* prevalence in small mammals (Insectivora and Rodentia). Folia Zoologica. 47(2): 93-101.

106. Herrmann, D.C.; Maksimov, P.; Maksimov, A.; Sutor, A.; Schwarz, S.; Jaschke, W.; Schliephake, A.; Denzin, N.; Conraths, F.J. and Schares, G. (2012). *Toxoplasma gondii* in foxes and rodents from the German Federal States of Brandenburg and Saxony - Anhalt: Seroprevalence and genotypes. Veterinary Parasitology. 185(2-4): 78 – 85.

107. Hide, G. (2016). Role of vertical transmission of *Toxoplasma gondii* in prevalence of infection. Expert Review of Anti-infective Therapy. 14(3): 335-44.

108. Hide, G.; Morley, E.K.; Hughes, J.M.; Gerwash, O.; Elmahaishi, M.S.; Elmahaishi, K.H.; Thomasson, D.; Wright, E.A.; Williams, R.H.; Murphy, R.G. and Smith, J.E. (2009). Evidence for high levels of vertical transmission in *Toxoplasma gondii*. Parasitology. 136(14): 1877-1885.

109. Higgs, S. and Nowell, F. (2000). Population biology of *Eimeria* (protozoa: Apicomplexa) in *Apodemus sylvaticus*: a capture study. Parasitology. 120(4): 355-363.

110. Homan, W.L.; Vercammen, M.; De Braekeleer, J.; Verschueren, H. (2000). Identification of a 200- to 300-fold repetitive 529 bp DNA fragment in *Toxoplasma gondii*, and its use for diagnostic and quantitative PCR. International Journal of Parasitology. 30(1): 69-75.

111. Hostomská, L.; Jírovec, O.; Horáčková, M.; Hrubcová, M. (1957) Účast toxoplasmické infekce matky při vniku mongoloidismu dítěte. (The role of toxoplasmosis in the mother in the development of mongolism in the child).Českoslovesnka Pediatrie 12:713–723.

112. Howe, D.K.; Sibley, L.D. (1995). *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. Journal of Infectious Diseases. 172(6): 1561–1566.

113. Howe, D.K.; Honore, S.; Derouin, F. and Sibley, L.D. (1997). Determination of genotypes of *Toxoplasma gondii* strains isolated from patients with toxoplasmosis. Journal of Clinical Microbiology. 35(6): 1411–1414.

114. Hökelek, M. (Updated 2013 May 3). Toxoplasmosis. Medscape Reference [On-line information]. Available online at http://emedicine.medscape.com/article/229969-overview through http://emedicine.medscape.com. Accessed September 2014.

115. Hughes, J.M.; Williams, R.H.; Morley, E.K.; Cook, D.A.N.; Terry, R.S.; Murphy, R.G.; Smith, J.E. and Hide, G. (2006). The Prevalence of Neospora *caninum* and co-infection with *Toxoplasma gondii* by PCR analysis in naturally occurring mammal populations. Parasitology 132(1): 29-36.

116. Hughes, J.M.; Thomasson, D.; Craig, P.S.; Georgin, S.; Pickles, A. and Hide, G. (2008). *Neospora caninum*: detection in wild rabbits and investigation of co-infection with *Toxoplasma gondii* by PCR analysis. Experimental Parasitology 120(3): 255-260.

117. Hurtado, A.; Aduriz, G.; Moreno, B.; Barandika, J.; García-Pérez, A.L. (2001). Single tube nested PCR for the detection of *Toxoplasma gondii* in foetal tissues from naturally aborted ewes. Veterinary Parasitology. 102(1-2): 17-27.

118. Hutchinson, W.M. (1965). Experimental transmission of *Toxoplasma gondii*. Nature. 206: 961–962.

119. Innes, E. A.; Bartley, P. M.; Buxton, D. and Katzer, F. (2009). Ovine toxoplasmosis. Parasitology. 136(14): 1887–1894.

120. Iqbal, J.; Khalid, N. (2007). Detection of acute *Toxoplasma gondii* infection in early pregnancy by IgG avidity and PCR analysis. Journal of medical Microbiology. 56(11): 1495-1499.

121. Jacobs, L.; Fair, J.R.; Bickerton, J.H (1954). Report of a Parasitologically Proved Case. AMA Archives of Ophthalmology. 52(1): 63-71.

122. Jackson, M.H.; Hutchison, W.M. and Siim, J.C. (1986). Toxoplasmosis in a wild rodent population of Central Scotland and a possible explanation of the mode of transmission. Journal of Zoology 209: 549-557.

123. Jacobs, L.; Remington, J.S.; Melton, M.L.; (1960). The resistance of the encysted form of *Toxoplasma gondii* Journal of Parasitology. 46: 11–21.

124. Janků, J. (1923). Pathogenesa a pathologická anatomie tak nazvaneho vrozeného kolombu žluté skvrany voku normálné velikem a microphthalmickém s nalezem parasitu v sítnici. Časopis lékařů českych. 62: 1021–1138.

125. Jeon, S.H. and Yong, T.S. (2000). Serological observation of *Toxoplasma gondii* prevalence in *Apodemus agrarius*, a dominant species of field rodents in Korea. Yonsei Medical Journ*al.* 41(4): 491-496.

126. Jewell, M.L.; Frenkel, J.K.; Johnson, K.M.; Reed, V.; Ruiz, A. (1972). Development of *Toxoplasma* oocysts in neotropical felidae. American Journal of Tropical Medicine and Hygiene. 21(5): 512–517.

127. Joiner, K.A. and Roos, D.S. (2002) Secretory traffic in *Toxoplasma gondii*: Less is more. Journal of Cell Biology. 157(4): 557-63.

128. Jones, J.L.; Kruszon-Moran, D. Wilson M. (2001). *Toxoplasma gondii* infection in the United States: seroprevalence and risk factors. American Journal of Epidemiology. 154(4): 357-365.

129. Jones, CD.; Okhravi, N.; Adamson, P.; Tasker, S.; Lightman, S. (2000). Comparison of PCR detection methods for B1, P30, and 18S rDNA genes of *T. gondii* in aqueous humour. Investigative Ophthalmology and Visual Science. 41(3): 634-644.

130. Jouquand, S.; Cheron, A.; Galibert, F. (1999). Microsatellite analysis using two-step procedure for fluorescent labelling of PCR products. Biotechniques. 26(5): 902-905.

131. Jung, C.; Lee, C.Y.; Grigg, M.E. (2004). The SRS superfamily of *Toxoplasma* surface proteins. International Journal of Parasitology. 34(3): 285-296.

132. Kang, H.K.; Cypel, Y.; Kilbourne, A.M.; Magruder, K.M.; Serpi, T.; Collins, J.F.; Frayne, S.M.; Furey, J.; Huang, G.D.; Kimerling, R.; Reinhard, M.J.; Schumacher, K.; Spiro, A. (2010). HealthViEWS: mortality study of female US Vietnam era veterans, 1965-2010. American Journal of Epidemiology. 179(6): 721-730.

133. Kasper, D.C.; Sadeghi, K, Prusa, AR.; Reischer, G.H.; Kratochwill, K.; Förster-Waldl, E.; Gerstl, N.; Hayde, M.; Pollak, A.; Herkner, K.R. (2009). Quantitative real-time polymerase chain reaction for the accurate detection of *Toxoplasma gondii* in amniotic fluid. Diagnostic Microbiology and Infectious Disease. 63(1): 10-5.

134. Khan, A.; Su, C.; German, M.; Storch, G.A.; Clifford, D.B. and Sibley, L.D. (2005). Genotyping of *Toxoplasma gondii* strains from immunocompromised patients reveals high prevalence of type I strains. Journal of Clinical Microbiology 43(12): 5881–5887.

135. Khan, A.; Dubey, J.P.; Su, C.; Ajioka, J.W.; Rosenthal, B.M.; Sibley, LD. (2011). Genetic analyses of atypical *Toxoplasma gondii* strains reveal a fourth clonal lineage in North America. International Journal of Parasitology. 41(6): 645-655.

136. Khemasuwan, D.; Mehta, A.C.; Wang, K.P. (2015). Past, present, and future of endobronchial laser photoresection. Journal of Thoracic Disease. 7(4): 380-388.

137. Kijlstra, A.; Meerburg, B.; Cornelissen, J.; De Craeye, S.; Vereijken, P. and Jongert, E. (2008). The role of rodents and shrews in the transmission of *Toxoplasma gondii* to pigs. Veterinary Parasitology. 156(3-4): 183-190.

138. Kim, J.O.; Jung, S.S.; Kim, S.Y.; Kim, T.Y.; Shin, D.W.; Lee, J.H.; Lee, Y.H. (2007). Inhibition of Lewis lung carcinoma growth by *Toxoplasma gondii* through induction of Th1 immune responses and inhibition of angiogenesis. Journal of Korean Medical Science. 22: 38-46.

139. Knani, L.; Bouslama, K.; Varette, C.; Gonzalez Canali, G.; Cabane, J.; Lebas, J.; Imbert, J.C. (1990). Pulmonary toxoplasmosis in AIDS. Report of 3 cases. [Case Reports, English Abstract, Journal Article]. Les Annales de Médecine Interne (Paris) .141(5): 469-471.

140. Kodym, P.; Machala, L.; Rohácová, H.; Sirocká, B.; Malý, M. (2007). Evaluation of a commercial IgE ELISA in comparison with IgA and IgM ELISAs, IgG avidity assay and complement fixation for the diagnosis of acute toxoplasmosis. Clinical Microbiology and Infection. 13(1): 40-47.

141. Kompalic-Cristo, A.; Frotta, C.; Suárez-Mutis, M.; Fernandes, O.; Britto, C. (2007). Evaluation of a real-time PCR assay based on the repetitive B1 gene for the detection of *Toxoplasma gondii* in human peripheral blood. Parasitology Research. 1(3): 619-625.

142. Krick, J. A.; Remington, J. S. (1978). Toxoplasmosis in adult – an overview. New England Journal of Medicine. 298(10): 550–553.

143. Kusbeci, O.Y.; Miman, O.; Yaman, M.; Aktepe, O.C.; Yazar, S. (2011). Could *Toxoplasma gondii* have any role in Alzheimer disease? Alzheimer Disease and Associated Disorders. 25(1): 1-3.

144. Lacy, R.C. (1987). Loss of Genetic Diversity from Managed Populations: Interacting Effects of Drift, Mutation, Immigration, Selection, and Population Subdivision. Conservation Biology. 1(2): 143–158.

145. La Vecchia, C.; Bosetti, C.; Lucchini, F.; Bertuccio, P.; Negri, E.; Boyle, P.; Levi, F. (2010). Cancer mortality in Europe, 2000-2004, and an overview of trends since 1975. Annals of Oncology. 21(6): 1323-1360.

146. Lass, A.; Pietkiweicz, H.; Modzelewska, E.; Dumètre, A.; Szostakowska, B.; Myjak , P. (2009). Detection of *Toxoplasma gondii* oocysts in environmental soil samples using molecular methods. European Journal of Clinical Microbiology and Infectious Diseases. 28(6): 599–605.

147. Lekutis, C.; Ferguson D.J.; Grigg M.E.; Camps M.; Boothroyd J.C. (2001). Surface antigens of *Toxoplasma gondii*: variations on a theme. International Journal of Parasitology. 31(12): 1285-1292.

148. Li, Y.C.; Korol, A.B.; Fahima, T.; Beiles, A.; Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology. 11 (12): 2453-2465.

149. Li, Z.; Zhao, Z.J.; Zhu, X.Q.; Ren, Q.S.; Nie, F.F.; Gao, J.M.; Gao, X.J.; Yang, T.B.; Zhou, W.L.; Shen, J.L.; Wang, Y.; Lu, F.L.; Chen, X.G.; Hide, G.; Ayala, F.J. and Lun, Z.R. (2012). Differences in iNOS and arginase expression and activity in the macrophages of rats are responsible for the resistance against *T. gondii* infection. PLoS One 7, e35834.

150. Liesenfeld, O.; Press, C.; Montoya, JG.; Gill, R.; Isaac-Renton, JL.; Hedman, K.; Remington, JS. (1997). False-positive results in immunoglobulin M (IgM) toxoplasma antibody tests and importance of confirmatory testing: the Platelia Toxo IgM test. Journal of Clinical Microbiology. 35(1): 174-178.

151. Lilue, J.; Muller, U.B.; Steinfeld, T. and Howard, J.C. (2013). Reciprocal virulence and resistance polymorphism in the relationship between *Toxoplasma gondii* and the house mouse. Elife 2, e01298

152. Ling, V.J.; Lester, D.; Mortensen, P.B.; Langenberg, P.W.; Postolache, T.T. (2011). *Toxoplasma gondii* seropositivity and suicide rates in women. Journal of Nervous and Mental Disease. 199(7): 440–444.

153. Loman, J. (1991). Do wood mice *Apodemus sylvaticus* abandon fields during autumn? Ekologia Polska 39:221-228.

154. Lopez, A.; Dietz, V.; Wilson, M.; Navin, T.; and Jones, J. (2000). Preventing congenital toxoplasmosis. MMWR. Recommendations and reports: Morbidity and

mortality weekly report. Recommendations and reports/Centres for Disease Control. 49 (2): 59-68.

155. Lundén, A.; Uggla, A. (1992) Infectivity of *Toxoplasma gondii* in mutton following curing, smoking, freezing or microwave cooking. International Journal of Food Microbiology; 15(3-4): 357-363.

156. Luft, B.J.; Brooks, R.G.; Conley, F.K.; McCabe, R.E.; Remington, J.S. (1984). Toxoplasmic encephalitis in patients with acquired immune deficiency syndrome. JAMA. 252(7): 913–917.

157. Luikart, G.; Sherwin, W.B.; Steele, BM.; Allendorf, FW. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology. 7(8): 963-974.

158. Makova, K.D.; Patton, J.C.; Krysanov, Chesser, R.K. and Baker, R.J. (1998). Microsatellite markers in wood mouse and striped field mouse (genus *Apodemus*). Molecular Ecology. 7: 247-249.

159. Marche, C.; Tabbara, W.; Michon, C.; Clair, B.; Bricaire, F.; Matthiessen, L. (1990). Bone marrow findings in HIV infection: a pathological study. Progress in AIDS pathology. 2: 51-60.

160. Marković, M.; Ivović, V.; Stajner, T.; Djokić, V.; Klun, I.; Bobić, B.; Nikolić, A.; Djurković-Djaković, O. (2014). Evidence for genetic diversity of *Toxoplasma gondii* in selected intermediate hosts in Serbia. Comparative Immunology, Microbiology and Infectious Diseases. 37(3): 173-179.

161. Marshall, P.A.; Hughes, J.M.; Williams, R.H.; Smith, J.E.; Murphy, R.G. and Hide, G. (2004). Detection of high levels of congenital transmission of *Toxoplasma gondii* in natural urban populations of *Mus domesticus*. Parasitology 128(1): 39-42.

162. Maruyama, T.; Fuerst, PA. (1985). Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. Genetics. 111(3): 675-689.

163. Meerburg, B.G.; De Craeye, S.; Dierick, K. and Kijlstra, A. (2012). *Neospora caninum* and *Toxoplasma gondii* in brain tissue of feral rodents and insectivores caught on farms in the Netherlands. Veterinary Parasitology. 184(2-4): 317-320.

164. Mercier, A.; Garba, M.; Bonnabau, H.; Kane, M.; Rossi, J.P.; Dardé, M.L.; Dobigny, G. (2013). Toxoplasmosis seroprevalence in urban rodents: a survey in Niamey, Niger. Memórias do Instituto Oswaldo Cruz. 108(4): 399-407.

165. Messaritakis, I.; Detsika, M.; Koliou, M.; Sifakis, S.; Antoniou, M. (2008). Prevalent genotypes of *Toxoplasma gondii* in pregnant women and patients from Crete and Cyprus. American Journal of Tropical Medicine and Hygiene. 79(2): 205-209.

166. Messier, V.; Lévesque, B.; Proulx, J.F.; Rochette, L.; Libman, M.D.; Ward, B.J.; Serhir, B.; Couillard, M.; Ogden, N.H.; Dewailly, E.; Hubert, B.; Déry, S.; Barthe, C.; Murphy, D.; Dixon, B. (2009). Seroprevalence of *Toxoplasma gondii* among Nunavik Inuit (Canada). Zoonoses Public Health. 56(4): 188-97.

167. Miller, M.A.; Miller, W.A.; Conrad, P.A.; James, E.R.; Melli, A.C.; Leutenegger, C.M.; Dabritz, H.A.; Packham, A.E.; Paradies, D.; Harris, M.; Ames, J.; Jessup, D.A.; Worcester, K.; Grigg, M.E. (2008). Type X *Toxoplasma gondii* in a wild mussel and terrestrial carnivores from coastal California: new linkages between terrestrial mammals, runoff and toxoplasmosis of sea otters. International Journal of Parasitology. 38(11): 1319-1328.

168. Miller, N.L.; Frenkel, J.K.; Dubey, J.P. (1972). Oral infections with *Toxoplasma* cysts and oocysts in felines, other mammals, and in birds. Journal of Parasitology. 58(5): 928–937.

169. Mills, J. (1986). *Pneumocystis carinii* and *Toxoplasma gondii* infections in patients with AIDS. Reviews of infectious diseases. 8(6): 1001-1011.

170. Miman, O.; Kusbeci, O.Y.; Aktepe, O.C.; Cetinkaya, Z. (2010). The probable relation between *Toxoplasma gondii* and Parkinson's disease. Neuroscience Letters. 475(3): 129–131.

171. Miyahara, K.; Yokoo, N.; Sakurai, H.; Igarashi, I.; Sakata, Y.; Yoshida, Y.; Saito, A.; Hirose, T.; Suzuki, N. (1992). Antitumor activity of *Toxoplasma* lysate antigen against methylcholanthrene-induced tumor-bearing rats. Journal of Veterinary Medical Science. 54(2): 221-228.

172. Montgomery, W.I. (1989). Population regulation in the wood mouse, *Apodemus sylvaticus.* II. Density dependence in spatial distribution and reproduction. Journal of Animal Ecology. 58(2): 477-494.

173. Montgomery, I.; Dowie, M. (1993). The distribution and population regulation of the wood mouse *Apodemus sylvaticus* on field boundaries of pastoral farmland. Journal of Applied Ecology. 30(4): 783-791.

174. Montoya, J.G.; Remington, J.S. (1996). Toxoplasmic chorioretinitis in the setting of acute acquired toxoplasmosis. Clinical Infectious Diseases. 23(2): 277–282.

175. Montoya, JG.; Liesenfeld, O. (2004). Toxoplasmosis. Lancet. 363(9425):1965-76.

176. Morley, E.K.; Williams, R.H.; Hughes, J.M.; Terry, R.S.; Duncanson, P.; Smith, J.E. and Hide, G. (2005). Significant familial differences in the frequency of abortion and *Toxoplasma gondii* infection within a flock of Charollais sheep. Parasitology. 131(2): 181-185.

177. Morley, E.K.; Williams, R.H.; Hughes, J.M.; Thomasson, D.; Terry, R.S.; Duncanson, P.; Smith, J.E. and Hide, G. (2008). Evidence that primary infection of Charollais sheep with *Toxoplasma gondii* may not prevent foetal infection and abortion in subsequent lambings. Parasitology 135(2): 169-173.

178. Morsy, T.A.; El Bahrawy, A.F.; El Dakhil, M.A. (2001). Ecto- and blood parasites affecting *Meriones rex* trapped in Najran, Saudi Arabia. Journal of the Egyptian Society of Parasitology. 31(2): 399-405.

179. de Moura, L.; Bahia-Oliveira, L.M.; Wada, M.Y.; Jones, J.L.; Tuboi, S.H.; Carmo, E.H.; Ramalho, W.M.; Camargo, N.J.; Trevisan, R.; Graça, R.M.; da Silva, A.J.; Moura, I.; Dubey, J.P.; Garrett, D.O. (2006). Waterborne toxoplasmosis, Brazil, from field to gene. Emerging Infectious Disease journ*al.* 12(2): 326-9.

180. Murphy, R.G.; Williams, R.H.; Hughes, J.M.; Hide, G.; Ford, N.J. and Oldbury, D.J. (2008). The urban house mouse (*Mus domesticus*) as a reservoir of infection for the human parasite *Toxoplasma gondii*: an unrecognised public health issue? International Journal of Environmental Health Research 18(3): 177-185.

181. Nagel, S.D. and Boothroyd, J. C. (1989). The major surface antigen of *Toxoplasma gondii* is anchored by a glycolipid. Journal of Biological Chemistry. 264(10): 5569–5574.

182. Naot, Y.; Remington, J. S. (1980). An enzyme-linked immunosorbent assay for detection of IgM antibodies to *Toxoplasma gondii*: use for diagnosis of acute acquired toxoplasmosis. Journal of Infectious Diseases. 142(5): 757-766.

183. Nei, M.; Maruyama T.; Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. Evolution. 29(1): 1-10

184. Nei, M.; Li, W.H. (1976). The transient distribution of allele frequencies under mutation pressure. Genetics Research. 28(3): 205-14.

185. Nicolle, C.; Manceaux, L. (1909). Sur un protozoaire nouveau du gondi. CR Academy Science. 148: 369.

186. Niebuhr, D.W.; Millikan, A.M.; Cowan, D.N.; Yolken, R.; Li, Y.; Weber, N.S. (2008). Selected infectious agents and risk of schizophrenia among U.S. military personnel. American Journal of Psychiatry. 165(1) :99–106.

187. Nimir, A.; Othman, A.; Ee, S.; Musa, Z.; Majid, IA.; Kamarudin, Z.; Xian, C.; Isa, NH. (2010). Latent toxoplasmosis in patients with different malignancy: a hospital based study. Journal of Clinical Medicine and Research. 2(3): 117-120.

188. Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample. Ecological Applications. 1(3): 462-474.

189. Nowak, R.M. 1991. Walker's Mammals of the World. 5th Ed. Vol II. Johns Hopkins University Press, Baltimore.

190. Nowakowska, D.; Colón, I.; Remington, J.S.; Grigg, M.; Golab, E.; Wilczynski, J.; Sibley, LD. (2006). Genotyping of *Toxoplasma gondii* by multiplex PCR and peptide-based serological testing of samples from infants in Poland diagnosed with congenital toxoplasmosis. Journal of Clinical Microbiology. 44(4): 1382-1389.

191. Olak, J.; Colson, Y. (2004). Gender differences in lung cancer: have we really come a long way, baby? Journal of Thoracic and Cardiovascular Surgery. 128(3): 346-51.

192. Overdulve, J.P. (1970). The identity of *Toxoplasma* Nicolle & Manceaux, 1909 with *Isospora Schneider*, 1881 (I). Proceedings Koninklijke Nederlandse Akademie der Wetenschappen, Series C. 73: 129–151.

193. Owen, M.R. and Trees, A.J. (1998). Vertical transmission of *Toxoplasma gondii* from chronically infected house (*Mus musculus*) and field (*Apodemus sylvaticus*) mice determined by polymerase chain reaction. Parasitology. 116(4): 299-304.

194. Parker, S.P.; ed. 1990. Grziemek's Encyclopedia. Vol 3. McGraw-Hill Publishers, NY.

195. Parkin, DM. (2010). Tobacco-attributable cancer burden in the UK in 2010. British Journal of Cancer. 105(2): 6-13.

196. Peakall, R.; Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 28(19): 2537-2539.

197. Pena, H.F.; Soares, R.M.; Amaku, M.; Dubey, J.P.; Gennari, S.M. (2006). *Toxoplasma gondii* infection in cats from São Paulo state, Brazil: seroprevalence, oocyst shedding, isolation in mice, and biologic and molecular characterization. Research in Veterinary Science. 81(1): 58-67.

198. Petersen, E.; Dubey, J.P. (2001). Biology of Toxoplasmosis. In Clinical Toxoplasmosis: Prevention and Management, D. H. M. Joynson and T. G. Wreghitt (Eds.). pp. 1-42. Cambridge University Press, Cambridge 2001.

199. Piry, S. Luikart, G.; Cornuet, JM. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. Journal of Heredity. 90(4): 502-503.

200. Piry, S.; Alapetite, A.; Cornuet, J.-M.; Paetkau, D.; Baudouin, L.; Estoup, A. (2004). GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. Journal of Heredity. 95(6): 536-539.

201. Pomeroy, C.; Filice, G.A. (1992). Pulmonary toxoplasmosis: a review. Clinical Infectious Diseases. 14(4): 863–870.

202. Porter, S.B.; Sande, M.A. (1992). Toxoplasmosis of the central nervous system in the acquired immunodeficiency syndrome. New England Journal of Medicine. 327(23): 1643-1648.

203. Prandovszky, E.; Gaskell, E.; Martin, H.; Dubey, J.P.; Webster, J.P.; McConkey, G.A. (2011). The neurotropic parasite *Toxoplasma gondii* increases dopamine metabolism. PLoS One. 6(9):e23866.

204. Pritchard, J.K.; Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155(2): 945–959.

205. Prestrud, K.W.; Dubey, J.P.; Asbakk, K.; Fuglei, E.; Su, C. (2008). First isolate of *Toxoplasma gondii* from arctic fox (*Vulpes lagopus*) from Svalbard. Veterinary Parasitology. 151(2-4):110-114.

206. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

207. Raizman, R.E. and Neva, F.A. (1975). Detection of circulating antigen in acute experimental infections with *Toxoplasma gondii*. Journal of Infectious Diseases.132: 44-48.

208. Raymond, M.; Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity. 86(3): 248-249

209. Röder, B.; Frühwirth, K.; Vogl, C.; Wagner, M.; Rossmanith, P. (2010). Impact of long-term storage on stability of standard DNA for nucleic acid-based methods. Journal of Clinical Microbiology. 48(11): 4260-4262.

210. Roe, W.D.; Howe, L.; Baker, E.J.; Burrows, L.; Hunter, S.A. (2013). An atypical genotype of *Toxoplasma gondii* as a cause of mortality in Hector's dolphins (*Cephalorhynchus hectori*). Veterinary Parasitology. 192(1-3): 67-74.

211. Roever-Bonnet, H. de (1961). Congenital toxoplasmosis. Tropical and geographical medicine. 13: 27-41.

212. Rodriguez, C.; Afchain, D.; Capron, A.; Dissous, C.; Santoro, F. (1985). Major surface protein of *Toxoplasma gondii* (p30) contains an immunodominant region with repetitive epitopesEuropean Journal of Immunology. 15(7): 747-749.

213. Rodriguez, K.A.; Guitron, J.; Hanseman, D.J.; Williams, V.; Starnes, S.L. (2012). Adjuvant chemotherapy and age-related biases in non-small cell lung cancer. Annals of Thoracic Surgery. 94(6): 1810-1814.

214. Roghmann M.C.; Faulkner C.T.; Lefkowitz A, Patton S, Zimmerman J, Morris J.G. (1999). Decreased seroprevalence for *Toxoplasma gondii* in Seventh Day Adventists in Maryland. American Journal of Tropical Medicine and Hygiene. 60: 790-792.

215. Rothova, A. (1993) Ocular involvement in toxoplasmosis. British Journal of Ophthalmology. 77(6): 371–377.

216. Ryan, P.; Hurley, S.; F.; Johnson, A.; M.; Salzberg, M.; Lee, M.; W.; North, J.; B.; McNeil, J.; J.; McMichael, A.; J. (1993). Tumours of the brain and presence of antibodies to *Toxoplasma gondii*. Internation*al*. Journal of Epidemiology. 22(3): 412–419.

217. Sasco, A.J.; Secretan, M.B.; Straif, K. (2004). Tobacco smoking and cancer: a brief review of recent epidemiological evidence. Lung Cancer. 45(2): 3-9.

218. Savva, D.; Morris, J.C.; Johnson, J.D. and Holliman, R.E. (1990). Polymerase chain reaction for detection of *Toxoplasma gondii*. Journal of Medical Microbiology. 32(1): 25–31.

219. Schaefer, J.J.; White, H.A.; Schaaf, S.L.; Mohammed, H.O.; Wade, S.E. (2011). Modification of a commercial *Toxoplasma gondii* immunoglobulin G enzyme-linked immunosorbent assay for use in multiple animal species. Journal of Veterinary Diagnostic Investigation. 23(2): 297-301.

220. Schneemann, M.; Schoedon, G.; Linscheid, P.; Walter, R.; Blau, N.; Schaffner, A. (1997). Nitrite generation in interleukin-4-treated human macrophage cultures does not involve the nitric oxide synthase pathway. Journal of Infectious Diseases. 175(1): 130-135.

221. Schreiber, R. D.; Feldman, H. A. (1980). Identification of the activator system for antibody to *Toxoplasma* as the classical complement pathway. Journal of Infectious Diseases. 141(3): 366–369.

222. Seaward, M.R.D.; Pentecost, A. (2001). Lichen flora of the Malham Tarn area. Field Studies. 10(1): 57 - 92.

223. Seyedeh, M.S.; Nahid, S.; Nahid, M.; Shima, D.P.; Morteza, Y.; Hossein, Y.D. (2015). Low titer of antibody against *Toxoplasma gondii* may be related to resistant to cancer. Journal of Cancer Research and Therapeutics. 11(2): 305-307.

224. Sheffield, H.G.; Melton, M.L. (1970). *Toxoplasma gondii*: the oocyst, sporozoite, and infection of cultured cells. Science. 167: 892–893.

225. Shen, Q.; Wang, L.; Fang, Q.; Shen, JL.; Zhang, LJ. (2014). Seroprevalance of *Toxoplasma gondii* infection and genotyping of the isolates from cancer patients in Anhui, eastern China. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 32(5): 366-70. Chinese.

226. Shwab, E.K.; Zhu, X.Q.; Majumdar, D.; Pena, H.F.; Gennari, S.M.; Dubey, J.P.; and Su, C. (2013). Geographical patterns of *Toxoplasma gondii* genetic diversity revealed by multilocus PCR-RFLP genotyping. Parasitology. 141(4): 453-461.

227. Sibley, L.D, Boothroyd, J.C. (1992). Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. Nature. 359(6390): 82-85.

228. Sinker, C.A. The vegetation of the Malham Tarn area. Leeds Philosophical and Literary Society, 1960.

229. Simpson, E.H. (1949). Measurement of Diversity. Nature. 163: 688-688.

230. Smith, D.D. and Frenkel, J.K. (1995). Prevalence of antibodies to *Toxoplasma gondii* in wild mammals of Missouri and East Central Kansas - biologic and ecologic considerations of transmission. Journal of Wildlife Diseases 31(1): 15-21.

231. Sommer, S.; Molecular Ecology Resources Primer Development Consortium. (2010). Permanent Genetic Resources added to the Molecular Ecology Resources Database 1 February 2010-31 March 2010. Molecular Ecology Resources. 10(4): 751-754.

232. Soulsby, E.J.L. (1978). Helminths, Arthropods and Protozoa of Domesticated Animals, 6th edition. The English Language Book Society and Bailliere Tindal and Cassall, London.

233. Spisak, F.; Turcekova, L.; Reiterova, K.; Spilovska, S. and Dubinsky, P. (2010). Prevalence estimation and genotypization *of Toxoplasma gondii* in goats. Biologia. 65(4): 670-674.

234. Splendore A. (1908). Un nuovo protozoa parassita de' conigli. incontrato nelle lesioni anatomiche d'une malattia che ricorda in molti punti il Kala-azar dell' uomo. Nota preliminare pel. Rev Soc Scient Sao Paulo. 3:109–112.

235. Stabile, L.P.; Davis, A.L.; Gubish, C.T.; Hopkins, T.M.; Luketich, J.D.; Christie, N.; Finkelstein, S.; Siegfried, J.M. (2002). Human non-small cell lung tumours and cells derived from normal lung express both oestrogen receptor alpha and beta and show biological responses to oestrogen. Cancer Research. 62(7): 2141-2150.

236. Stahl, W.; Sekiguchi, M. and Kaneda Y. (1998). Cerebellar anomalies in congenital murine toxoplasmosis. Tokai Journal of Experimental and Clinical Medicine. 23(6): 261-265.

237. Storz, J.F.; Bhat, H.R.; Kunz, T.H. (2001). Genetic consequences of polygyny and social structure in an Indian fruit bat, *Cynopterus sphinx*. I. Inbreeding, outbreeding, and population subdivision. Evolution. 55(6): 1215-23.

238. Su, C.; Zhang, X. and Dubey, J.P. (2006). Genotyping of *Toxoplasma gondii* by multilocus PCR-RFLP markers: a high resolution and simple method for identification of parasites. International Journal of Parasitology. 36: 841 – 848.

239. Sun, S.; Schiller, J.H.; Gazdar, A.F. (2007). Lung cancer in never smokers--a different disease. Nature Reviews Cancer. 7(10): 778-90.

240. Sutherland, G. R.; Richards, R. I. (1992). Anticipation legitimized: unstable DNA to the rescue. (Editorial) American Journal of Human Genetics. 51(1): 7-9.

241. Suzuki, Y.; Joh, K.; Orellana, M.A.; Conley, F.K.; Remington, J.S. (1991). A gene(s) within the H-2D region determines the development of toxoplasmic encephalitis in mice. Immunology. 74(4):732–739.

242. Suzuki, Y.; Orellana, M.A.; Wong, S.Y.; (1993). Susceptibility to chronic infection with *Toxoplasma gondii* does not correlate with susceptibility to acute infection in mice. Infection and Immunity. 61(6): 2284–2288.

243. Suzuki, Y.; Claflin, J.; Wang, X.; Lengi, A. and Kikuchi, T. (2005). Microglia and macrophages as innate producers of interferon-gamma in the brain following infection with *Toxoplasma gondii*. International Journal of Parasitology. 35(1): 83-90,

244. Tenter, A.M.; Heckeroth, A.R. and Weiss, L.M. (2000). *Toxoplasma gondii*: from animals to humans. International Journal for Parasitology. 30(12-13): 1217-1258.

245. Terry, R.S.; Smith, J.E.; Duncanson, P. and Hide, G. (2001). MGE-PCR: a novel approach to the analysis of *Toxoplasma gondii* strain differentiation using mobile genetic elements. International Journal for Parasitology. 31(2): 155-161.

246. Tew, T.E. and Macdonald, D.W. (1993). The effects of harvest on arable field mice, *Apodemus sylvaticus*. Biological Conservation 65(3): 279-283.

247. Thiermann, E. (1957). Transmission congenita del *Toxoplasma gondii* en ratas con infeccion leve. Biologica 23: 59.

248. Thirugnanam, S.; Rout, N.; Gnanaseka, M. (2013). Possible role of *Toxoplasma gondii* in brain cancer through modulation of host microRNAs. Infectious Agents and Cancer. 8:8.

249. Thomas, F.; Lafferty, K.D.; Brodeur, J.; Elguero, E.; Gauthier-Clerc, M.; Misse, D. (2012). Incidence of adult brain cancers is higher in countries where the protozoan parasite *Toxoplasma gondii* is common. Biology Letters. 8: 101-103.

250. Thomasson, D.; Wright, E. A.; Hughes, J. M.; Dodd, N.S.; Cox, A. P.; Boyce, K.; Gerwash, O.; Abushahma, M.; Lun, Z. R.; Murphy, R.G.; Rogan, M. T. and Hide. G. (2011). Prevalence and Co-infection of *Toxoplasma gondii* and *Neospora caninum* in *Apodemus sylvaticus* in an area relatively free of cats. Parasitology. 138(9): 1117-1123.

251. Tibayrenc, M.; Kjellberg, F.; Ayala, F.J. (1990). A clonal theory of parasitic protozoa: the population structure of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas* and *Trypanosoma,* and its medical and taxonomical consequences. Proceedings of the National Academy of Sciences USA. 87(7): 2414–2418.

252. Van Oosterhout, C.; Hutchinson, W.; Wills, D. and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes. 4(3): 535–538.

253. Varilo, T.; Paunio, T.; Parker, A.; Perola, M.; Meyer, J.; Terwilliger, J.D.; Peltonen, L. (2003). The interval of linkage disequilibrium (LD) detected with microsatellites and SNP markers in chromosomes of Finnish populations with different histories. Human Molecular Genetics. 12(1): 51-59.

254. Vestbo, J. (2013). COPD: definition and phenotypes. Clinics in Chest Medicine. 35(1): 1-6.

255. Vidal, J. E.; Colombo, F.; A.; Penalva de Oliveira, A. C.; Focaccia, R.; Pereira-Chioccola, V. L. (2004). PCR assay using cerebrospinal fluid for diagnosis of cerebral toxoplasmosis in Brazilian AIDS patients. Journal of Clinical Microbiology. 42(10): 4765-4768.

256. Villena, I.; Marle, M.; Dardé, M.L.; Pinon, J.M.; Aubert, D. (2004). *Toxoplasma* strain type and human disease: risk of bias during parasite isolation? Trends in Parasitology. 20(4):160-162.

257. Vittecoq, M.; Elguero, E.; Lafferty, K.D.; Roche, B.; Brodeur J.; Gauthier-Clerc, M.; Missé D.; Thomas, F. (2012). Brain cancer mortality rates increase with *Toxoplasma gondii* seroprevalence in France. Infection, Genetics and Evolution. 12(2): 496-498.

258. Viard, F.; Bremond, P.; Labbo, R.; Justy, F.; Delay, B. and Jarne, P. (1996). Microsatellites and genetics of highly selfing populations in the freshwater snail *Bulinus tuncatus*. Genetics. 142(4): 1237-1247.

259. Vyas, A. (2013). Parasite-augmented mate choice and reduction in innate fear in rats infected by *Toxoplasma gondii*. Journal of Experimental Biology. 216(1):120-126.

260. Vujanić, M.; Ivović, V.; Kataranovski, M.; Nikolić, A.; Bobić, B.; Klun, I.; Villena, I.; Kataranovski, D. and Djurković-Djaković, O. (2012). Toxoplasmosis in naturally infected rodents in Belgrade, Serbia. Vector Borne Zoonotic Diseases. 11(8): 1209-1211.

261. Waap, H.; Vilares, A.; Rebelo, E.; Gomes, S.; Angelo, H. (2008). Epidemiological and genetic characterization of *Toxoplasma gondii* in urban pigeons from the area of Lisbon (Portugal). Veterinary Parasitology. 157(3-4): 306-309.

262. Wallace G.D. (1972). Experimental transmission of *Toxoplasma gondii* by cockroaches. Journal of Infectious Diseases. 126(5): 545–547.

263. Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. Genetics. 166(4): 1963-1979.

264. Watts, C.H.S. (1969). The regulation of wood mouse (*Apodemus sylvaticus*) numbers in Wytham Woods, Berkshire. Journal of Animal Ecology. 38(1):285– 304.

265. Weiland, G.; Kühn, D. (1970). Experimentelle *Toxoplasma*-Infektionen bei der Katze. II. Entwicklungsstadien des Parasiten im Darm. Berlin München Tierärztliche Wochenschrift. 83: 128–132.

266. Weller, P. Parasitic pneumonias. (1988). In: Pennington JE ed. Respiratory Infections. Diagnosis and Management, 2nd ed. New York: Raven Press.

267. Williams, R.H.; Morley, E.K.; Hughes, J.M.; Duncanson, P.; Terry, R.S.; Smith, J.E. and Hide, G. (2005). High levels of congenital transmission of *Toxoplasma gondii* in longitudinal and cross-sectional studies on sheep farms provide evidence of vertical transmission in ovine hosts. Parasitology. 130(3): 301-307.

268. Wilson W.L.; Montgomery, W.I.; Elwood, R.W. (1993). Population regulation in the wood mouse *Apodemus sylvaticus*. Mammal Review. 23: 73-92.

269. Witte, H.M.; Piekarski, G. (1970). Die Oocysten-Ausscheidung bei experimentell infizierten Katzen in Abhängigkeit vom *Toxoplasma*-Stamm. Zeitschrift fur Parasitenkunde. 33(4): 358–360.

270. Whitlock, M.C.; McCauley, D.E. (1999). Indirect measures of gene flow and migration: F_{ST} not equal to 1/(4Nm + 1). Heredity. 82(2): 117-125.

271. Wrensch, M.; Bondy, M.L.; Wiencke, J.; Yost, M. (1993). Environmental risk factors for primary malignant brain tumors: a review. The Journal of Neuro-Oncology. 17(1): 47-64.

272. Wrensch, M.; Minn, Y.; Chew, T.; Bondy, M.; Berger, M.; S. (2002). Epidemiology of primary brain tumors: current concepts and review of the literature. Neuro-Oncology. 4(4): 278–299.

273. Wolf, A.; Cowen, D.; Paige B. (1939). Human toxoplasmosis: occurrence in infants as encephalomyelitis verification by transmission to animals. Science. 89(2306): 226–227.

274. Wong, S.Y.; Remington, J.S. (1993). Biology of *Toxoplasma gondii*. AIDS. 7(3): 299-316.

275. Yazar, S.; Arman, F.; Yalcin, S.; Demirtas, F.; Yaman, O.; Sahin, I. (2003). Investigation of probable relationship between *Toxoplasma gondii* and cryptogenic epilepsy. Seizure. 12(2): 107–109.

276. Yazar, S.; Yaman, O.; Eser, B.; Altuntas, F.; Kurnaz, F.; Sahin, I. (2004). Investigation of anti-*Toxoplasma gondii* antibodies in patients with neoplasia. Journal of Medical Microbiology. 53(12): 1183-1186.

277. Yin, C.C.; He, Y.; Zhou, D.H.; Yan, C.; He, X.H.; Wu, S.M.; Zhou, Y.; Yuan, Z.G.; Lin, R.Q. and Zhu, X.Q. (2010). Seroprevalence of *Toxoplasma gondii* in rats in southern China. Journal of Parasitology. 96(3): 1233-1234.

278. Yolken, R.H.; Dickerson, F.B.; Fuller Torrey, E. (2009). *Toxoplasma* and schizophrenia. Parasite Immunology. 31(11): 706–715.

279. Yuan, Z.; Gao, S.; Liu, Q.; Xia, X.; Liu, X.; Liu, B.; Hu, R. (2007). *Toxoplasma gondii* antibodies in cancer patients. Cancer Letters. 254(1): 71–74.

280. Zenner, L.; Foulet, A.; Caudrelier, Y.; Darcy, F.; Gosselin, B.; Capron, A.; Cesbron-Delauw, MF. (1999). Infection with *Toxoplasma gondii* RH and Prugniaud strains in mice, rats and nude rats: kinetics of infection in blood and tissues related to pathology in acute and chronic infection. Pathology, Research and Practice. 95(7): 475-485.

281. Zhang, S.Y.; Jiang, S.F.; He, Y.Y.; Pan, C.E.; Zhu, M. and Wei, M.X. (2004). Serologic prevalence of *Toxoplasma gondii* in field mice, *Microtus fortis*, from Yuanjiang, Hunan province, People's Republic of China. Journal of Parasitology. 90(2): 437-438.

282. Zhao, Z.J.; Zhang, J.; Wei, J.; Li, Z.; Wang, T.; Yi, S.Q.; Shen, J.L.; Yang, T.B.; Hide, G. and Lun Z.R. (2013). Lower expression of inducible nitric oxide synthase

and higher expression of arginase in rat alveolar macrophages are linked to their susceptibility to *Toxoplasma gondii* infection. PLoS One 8, e63650.

283. Wang, S.; Lan, C.; Zhang, L.; Zhang, H.; Yao, Z.; Wang, D.; Ma, J.; Deng, J.; Liu, S. (2015). Seroprevalence of *Toxoplasma gondii* infection among patients with hand, foot and mouth disease in Henan, China: a hospital-based study. Infectious Diseases of Poverty. 4: 53.

284. Zubaid, A.; Gorman, M. L. (1993). The effect of supplementary feeding upon the demography of a population of wood mice *Apodemus sylvaticus*, living on a system of maritime sand-dunes. Journal of Zoology. 230: 609–617.

285. http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/ (accessed January 2016).

286. http://www.who.int/mediacentre/factsheets/fs315/en/ (accessed January 2016).

Appendix

Apodemus sylvaticus genotyping and population data (-9 represents missing value)

Expected size (bp)																				
	101-110		105-124		128-154		361-622		168-209		94-118		199-296		202-216		252-268			
Primer	8S		TNF		4A		3A		1A		2A		7S		9A		12A		MS19	
Sample	All 1	All 2	All 1	All 2	All 1	All 2	All 1	All 2	All 1	All 2	All1	All2	All1	All2	All1	All2	All1	All2	All1	All2
1	120	120	109	109	-9	-9	363	382	217	234	99	99	208	243	203	203	441	441	82	89
2	112	120	112	119	-9	-9	374	366	162	209	95	103	208	243	200	204	441	475	95	201
3	120	120	115	120	130	144	374	382	170	217	99	99	202	202	200	203	441	441	118	201
4	120	120	107	109	-9	-9	382	390	176	217	99	99	243	243	203	225	461	477	88	95
5	112	120	119	122	130	144	375	385	158	176	99	99	208	237	203	225	441	441	118	201
6	117	120	112	122	137	148	363	385	170	217	95	95	208	208	203	203	441	477	95	201
7	120	120	107	109	128	134	363	382	162	217	99	99	202	243	203	203	445	461	93	95
8	120	120	121	124	127	151	366	385	162	234	99	99	202	202	203	203	441	461	95	102
9	120	120	112	114	128	134	366	375	176	217	99	99	208	208	225	225	445	481	95	201
10	120	120	109	114	130	140	366	375	176	176	103	103	237	243	176	200	461	475	95	201
11	120	120	109	115	130	140	375	382	217	217	95	99	210	243	96	96	445	461	95	201
12	117	120	118	118	128	130	378	382	176	217	95	99	199	243	203	203	471	477	104	201
13	128	141	109	122	-9	-9	-9	-9	158	217	-9	-9	205	207	225	225	-9	-9	95	95
14	120	120	117	120	137	148	366	382	170	217	97	97	206	237	203	203	461	461	102	201
15	120	120	120	122	137	148	363	378	217	217	99	99	206	243	203	225	461	467	95	95
16	120	120	118	120	140	144	367	382	162	217	99	99	208	208	225	225	452	477	95	120
17	120	120	118	122	130	133	367	382	176	176	97	97	210	210	176	225	441	445	97	104
18	120	120	115	122	130	133	378	382	162	162	97	97	210	210	203	225	441	441	102	104
19	117	120	120	122	133	154	382	390	162	217	99	99	210	210	203	225	461	467	95	97
20	120	120	107	109	128	130	370	386	196	217	99	99	206	206	225	225	452	477	124	201
21	112	120	108	122	137	148	365	378	176	211	99	99	206	238	176	225	-9	-9	104	201
22	120	123	108	120	133	140	366	382	162	217	99	99	210	210	96	96	461	461	88	97
23	112	120	118	122	140	148	375	378	176	217	99	99	210	210	176	200	441	467	95	104
24	112	120	118	120	134	140	366	382	176	217	95	95	199	199	200	204	441	467	95	104
25	117	117	117	122	128	144	375	382	217	217	95	95	198	209	203	203	477	477	95	118
26	120	120	109	122	137	148	375	382	162	217	99	99	199	199	225	225	447	461	118	201
27	120	120	107	109	133	152	363	386	162	234	99	99	208	208	203	203	445	461	95	201
28	120	120	114	122	128	137	382	386	217	234	99	99	202	202	225	225	446	461	88	118

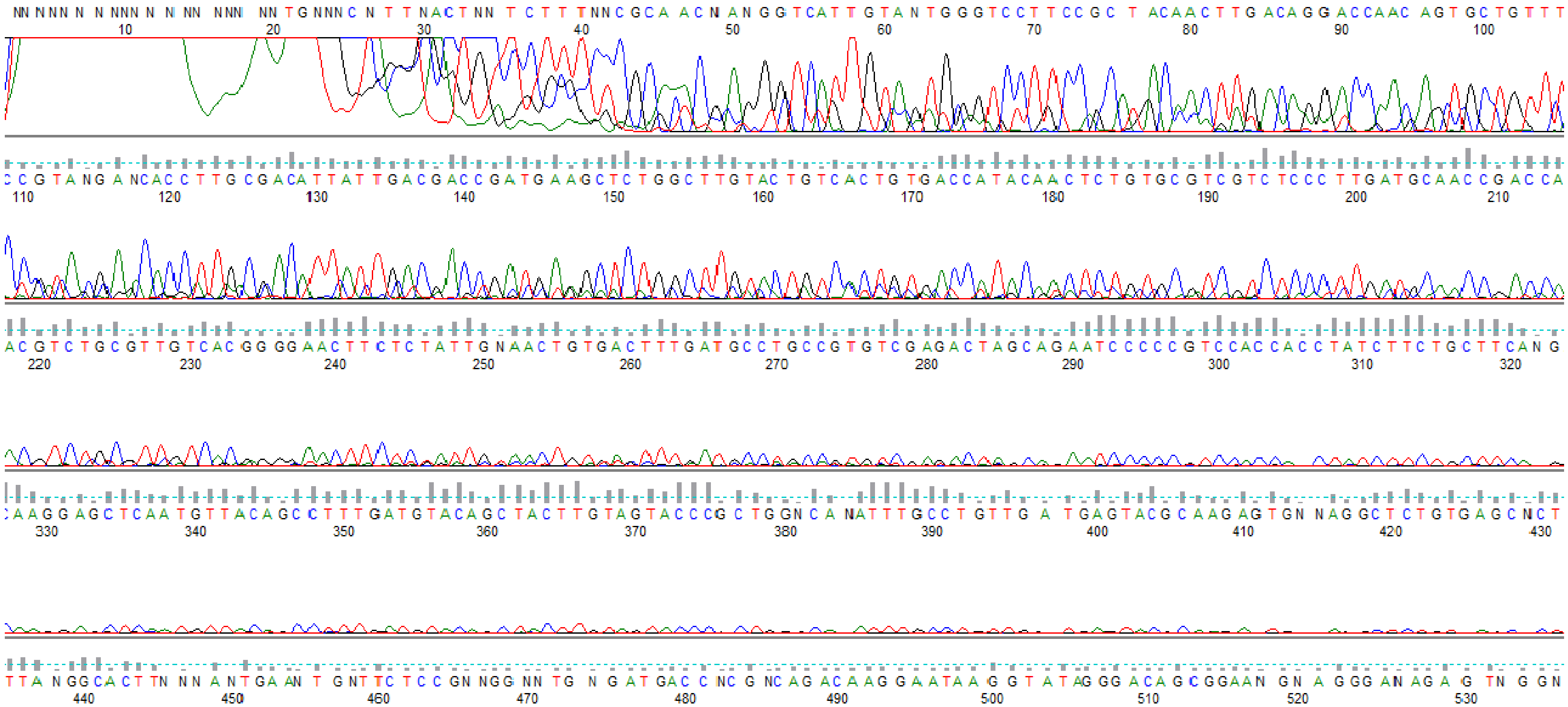
29	120	120	120	122	148	154	363	383	176	217	97	97	202	210	203	203	447	466	88	95
30	120	120	106	124	134	137	378	390	162	217	95	103	210	210	225	225	461	461	88	95
31	120	120	119	122	134	154	378	382	162	162	99	99	206	210	203	203	461	461	-9	-9
32	112	120	115	119	130	154	367	374	158	162	99	99	208	243	203	203	437	461	92	95
33	120	120	131	133	-9	-9	-9	-9	217	217	-9	-9	206	243	203	203	-9	-9	95	120
34	120	120	118	133	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	203	203	437	437	96	96
35	-9	-9	118	133	-9	-9	-9	-9	176	217	95	106	206	210	176	203	-9	-9	82	82
36	131	134	122	133	-9	-9	-9	-9	158	176	-9	-9	204	206	-9	-9	-9	-9	82	96
37	118	120	122	133	-9	-9	-9	-9	-9	-9	-9	-9	204	206	-9	-9	-9	-9	82	82
38	120	120	119	128	133	148	374	394	176	217	99	99	206	206	203	203	-9	-9	97	201
39	120	120	-9	-9	-9	-9	440	446	-9	-9	-9	-9	-9	-9	-9	-9	-9	-9	89	89
40	118	120	109	133	134	141	436	460	-9	-9	82	95	210	237	-9	-9	-9	-9	82	96
41	118	120	109	133	-9	-9	446	460	176	234	-9	-9	209	237	-9	-9	436	474	82	96
42	120	120	128	135	133	140	374	390	162	217	99	99	210	210	203	203	437	481	118	120
43	120	120	115	117	137	148	367	371	162	217	99	99	208	208	203	203	440	474	95	109
44	117	120	115	122	137	152	367	371	176	217	-9	-9	206	243	225	225	441	441	95	201
45	112	120	118	122	148	152	382	386	-9	-9	99	99	237	243	225	225	460	460	95	201
46	120	120	109	122	127	134	382	386	155	168	99	99	237	242	203	203	446	454	93	95
47	120	120	107	122	134	144	365	388	-9	-9	99	99	208	237	189	210	444	454	201	201
48	120	120	107	122	128	130	367	370	182	211	95	99	206	238	203	203	474	474	102	104
49	117	120	120	120	130	137	375	382	176	217	99	99	210	210	174	174	474	474	95	114
50	120	120	120	122	128	130	378	386	162	217	99	99	206	208	203	203	438	440	102	104
51	120	120	120	122	128	137	386	386	217	217	99	99	208	237	203	203	446	454	82	82
52	120	120	118	120	128	148	376	386	162	217	99	99	199	242	203	203	440	454	85	92
53	120	120	120	122	128	134	370	378	162	217	99	99	237	234	176	203	438	440	92	112
54	120	120	122	135	128	130	367	386	162	217	99	99	206	237	203	203	470	470	92	120
55	120	120	122	133	130	144	370	376	162	217	99	99	210	210	203	225	436	474	92	100
56	120	120	110	122	130	144	376	386	162	217	99	99	208	237	203	232	461	461	82	99
57	120	120	119	122	128	152	378	386	162	217	99	99	210	237	210	225	461	461	92	100
58	120	120	117	119	134	134	363	378	162	217	99	99	206	206	203	203	440	476	-9	-9
59	120	120	113	115	128	134	374	382	162	217	99	99	206	206	203	203	444	446	92	100
60	120	120	118	120	128	148	374	390	162	217	99	99	206	237	119	126	446	454	92	100
61	120	120	112	115	137	154	378	396	162	217	99	99	210	210	203	203	440	461	92	338
62	117	120	115	120	134	134	378	390	162	217	99	99	199	208	225	225	-9	-9	92	100

63	120	120	115	122	128	148	374	382	162	217	99	99	210	210	203	203	-9	-9	92	100
64	120	120	118	122	128	144	367	386	162	217	95	99	199	237	210	225	446	454	92	100
65	120	120	120	120	127	127	378	390	162	217	95	99	225	243	203	203	440	461	92	100
66	120	120	120	122	140	152	378	382	162	217	99	99	204	237	171	200	-9	-9	92	100
67	109	120	109	132	133	140	375	382	162	217	99	99	199	206	96	96	-9	-9	92	100
68	113	120	122	140	131	155	-9	-9	158	158	-9	-9	206	247	171	203	436	474	93	95
69	120	120	120	120	133	152	378	382	217	217	85	88	206	210	-9	-9	444	470	92	134
70	120	120	109	118	128	134	374	382	170	234	92	94	206	243	171	200	461	461	86	97
71	120	120	109	118	128	145	382	386	217	217	92	94	208	210	-9	-9	444	474	92	92
72	-9	-9	126	128	140	155	382	386	176	217	96	99	-9	-9	171	203	440	480	97	109
73	120	120	115	124	128	134	371	382	158	158	113	94	206	247	203	203	461	461	201	201
74	-9	-9	109	118	141	152	378	382	-9	-9	84	95	-9	-9	203	210	461	461	118	118
75	120	120	115	118	134	141	371	386	170	234	96	99	-9	-9	204	204	444	478	104	252
76	120	120	-9	-9	141	152	374	378	217	217	-9	-9	206	238	171	200	444	463	93	95
77	120	120	-9	-9	140	140	371	386	158	176	96	99	206	210	171	203	-9	-9	100	324
78	120	120	122	127	138	145	376	386	176	217	83	99	206	243	176	176	440	480	-9	-9
79	-9	-9	119	119	133	141	-9	-9	204	217	81	99	-9	-9	203	203	-9	-9	102	104
80	120	120	117	127	128	140	379	390	217	217	83	94	204	206	-9	-9	434	478	118	120
81	120	120	109	127	128	134	374	382	217	234	83	83	208	243	171	224	-9	-9	88	110
82	117	120	115	122	128	134	378	382	176	217	92	94	206	206	-9	-9	434	446	118	120
83	120	120	115	122	134	141	371	382	176	217	92	94	206	206	-9	-9	461	461	118	118
84	120	120	115	122	128	148	382	386	156	217	92	94	206	206	203	224	436	461	100	103
85	120	120	110	133	128	131	367	382	217	217	80	94	204	206	-9	-9	444	478	89	89
86	120	120	-9	-9	134	144	367	382	162	176	100	102	200	243	-9	-9	441	441	95	120
87	120	120	92	92	134	134	371	382	162	217	86	98	206	208	203	203	-9	-9	95	95
88	120	120	106	127	133	133	382	386	158	170	96	98	200	206	-9	-9	-9	-9	202	218
89	120	120	109	115	128	128	374	382	217	217	84	94	196	243	203	225	441	441	-9	-9
90	120	120	107	127	128	148	374	374	159	212	80	84	208	243	171	227	-9	-9	95	102
91	120	120	115	115	144	144	371	371	162	217	96	98	208	208	203	203	441	461	88	95
92	117	120	110	118	131	152	367	386	217	217	94	98	206	210	-9	-9	-9	-9	92	96
93	120	120	122	128	131	144	367	367	178	217	84	86	208	243	210	225	438	440	104	120
94	117	120	110	110	-9	-9	374	386	158	162	96	98	243	243	-9	-9	474	478	88	120
95	117	120	122	126	134	140	382	390	217	217	84	86	206	238	171	171	461	478	95	95
96	112	117	107	118	140	152	367	390	217	217	86	98	204	206	171	225	440	461	96	120

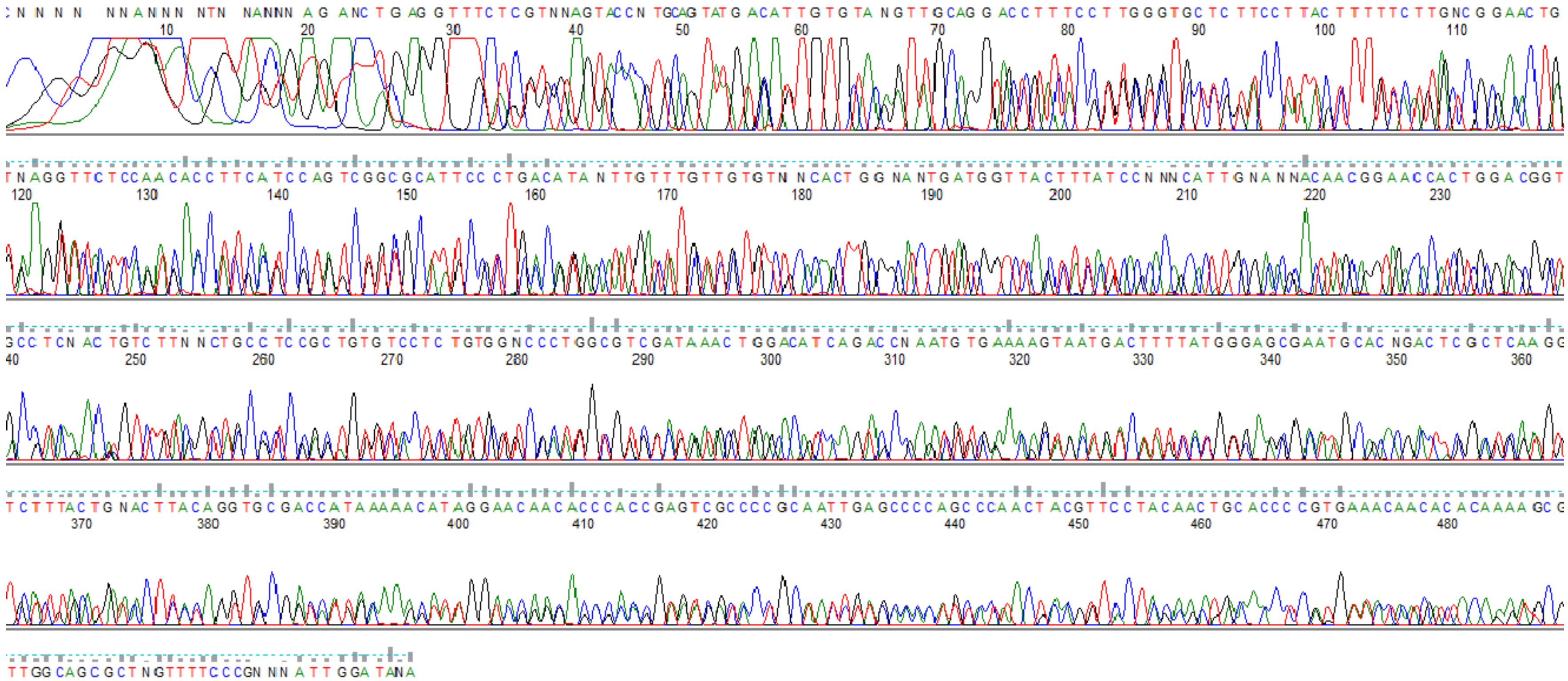
97	120	120	109	122	134	140	374	386	158	217	84	94	206	243	174	174	-9	-9	88	102
98	120	120	115	133	128	134	371	374	158	217	84	84	196	208	171	203	444	454	95	104
99	120	120	115	117	134	134	367	382	162	217	84	102	200	208	222	225	446	461	95	120
100	120	120	118	127	131	152	378	386	158	162	80	84	200	204	200	203	454	461	98	104
101	120	120	117	130	145	148	406	406	217	217	84	84	-9	-9	203	203	440	474	88	110
102	120	120	115	133	-9	-9	-9	-9	217	217	84	84	209	243	203	203	440	444	95	104
103	120	120	122	131	-9	-9	-9	-9	-9	-9	96	99	200	206	222	225	446	479	118	120
104	120	120	122	131	127	148	-9	-9	162	217	96	99	-9	-9	200	203	440	446	82	88
105	112	120	115	127	-9	-9	370	378	217	234	92	96	206	210	203	203	444	454	82	82
106	118	120	112	115	117	127	370	378	162	176	92	96	202	206	203	203	416	444	96	96
107	118	120	124	127	-9	-9	371	382	162	162	120	120	206	210	203	203	430	461	96	96
108	118	120	124	139	-9	-9	-9	-9	217	217	120	120	206	210	203	224	440	461	96	104
109	118	120	124	127	-9	-9	-9	-9	168	217	120	120	209	243	203	203	436	461	104	404
110	118	120	106	127	147	154	-9	-9	162	217	120	120	202	206	169	169	-9	-9	96	96
111	120	120	106	127	-9	-9	370	378	216	234	112	120	204	206	203	203	461	470	82	96
112	101	120	122	126	128	133	367	378	162	176	118	120	198	200	203	224	436	461	82	82
113	101	120	109	133	-9	-9	371	384	162	162	118	120	198	200	171	186	454	454	89	89
114	101	120	115	127	128	133	-9	-9	162	162	118	120	200	206	171	186	-9	-9	96	120
115	101	120	106	127	133	144	378	382	217	217	118	120	206	210	171	225	-9	-9	82	82
116	113	120	116	119	81	81	370	370	158	176	118	120	202	206	169	203	-9	-9	82	82
117	113	120	-9	-9	127	144	-9	-9	162	217	86	96	206	237	171	203	440	440	92	92
118	101	120	-9	-9	81	81	-9	-9	-9	-9	101	120	235	237	169	203	-9	-9	85	97
119	118	120	-9	-9	137	141	-9	-9	217	234	101	120	229	243	225	225	440	450	-9	-9
120	100	120	122	127	81	81	371	386	162	217	101	120	208	243	171	225	436	461	93	95
121	113	120	113	133	144	148	382	382	176	217	101	120	205	243	171	186	429	461	125	125
122	120	120	113	119	140	144	363	367	176	196	113	120	206	243	171	186	429	461	108	110
123	100	120	115	122	144	154	176	217	176	216	89	99	210	237	171	186	446	472	-9	-9
124	120	120	115	139	127	133	374	382	-9	-9	101	120	-9	-9	171	203	446	446	82	120
125	-9	-9	-9	-9	127	133	-9	-9	-9	-9	118	120	204	242	171	186	-9	-9	82	96
126	100	120	126	127	-9	-9	370	382	176	217	100	120	202	243	169	183	-9	-9	82	110

T. gondii sequence data obtained from the lung cancer patients

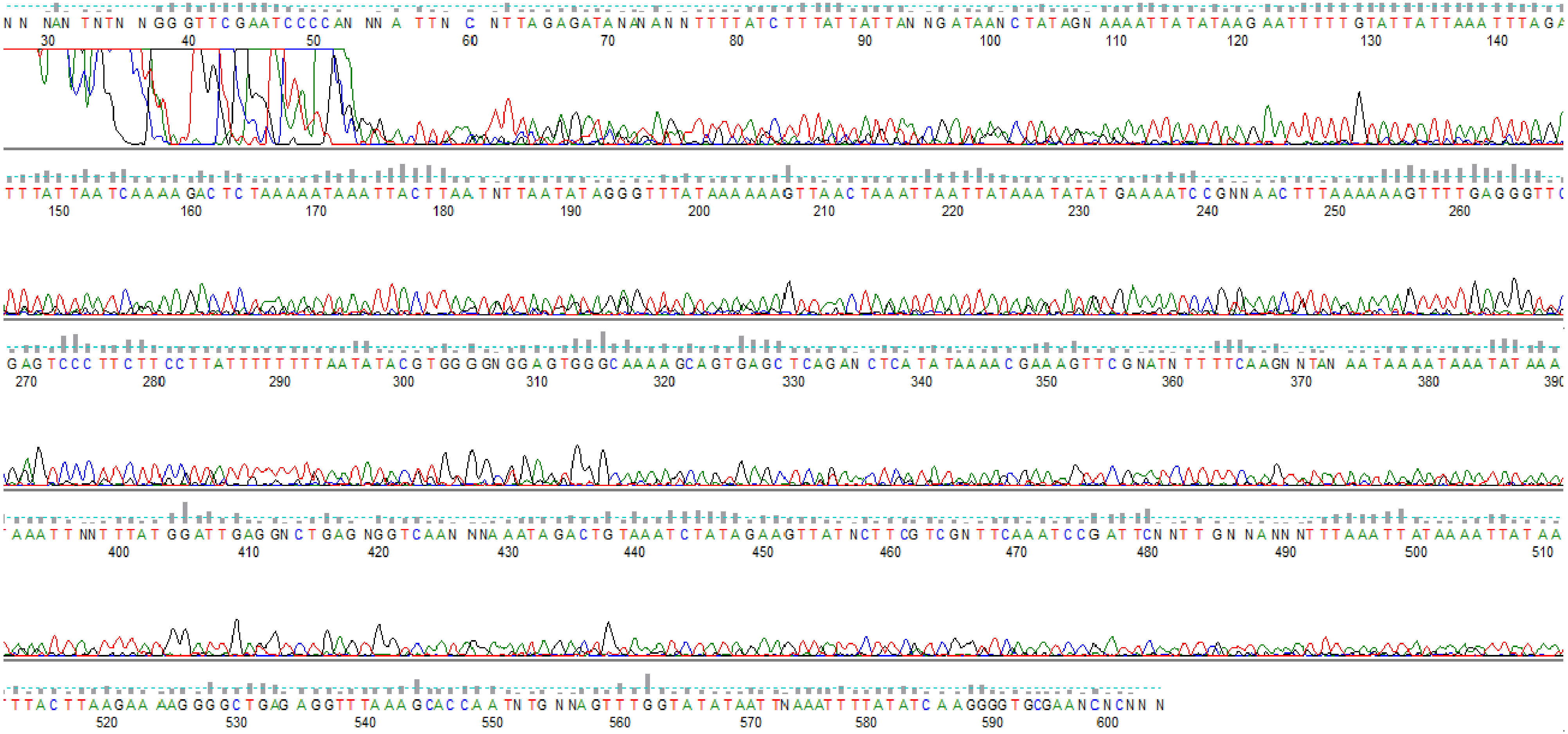
Sequence analysis of SAG1 marker (Savva *et al.* 1990) (subject 1051) was identified by BLAST as *T. gondii* VEG strain (95%).



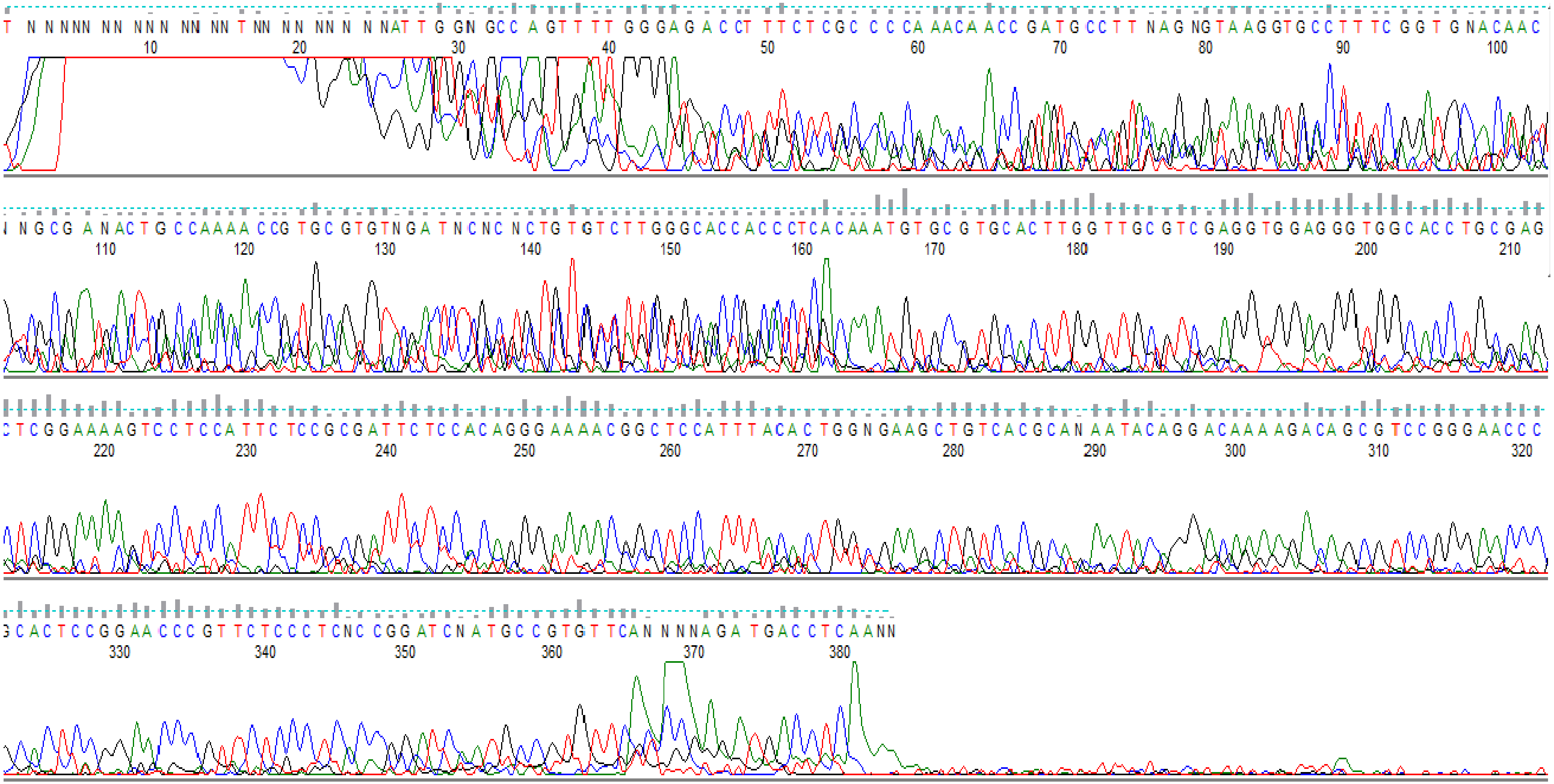
Sequence analysis of the alt.SAG2 marker (subject 918) was identified by BLAST as *T. gondii* ME49 strain (93%).



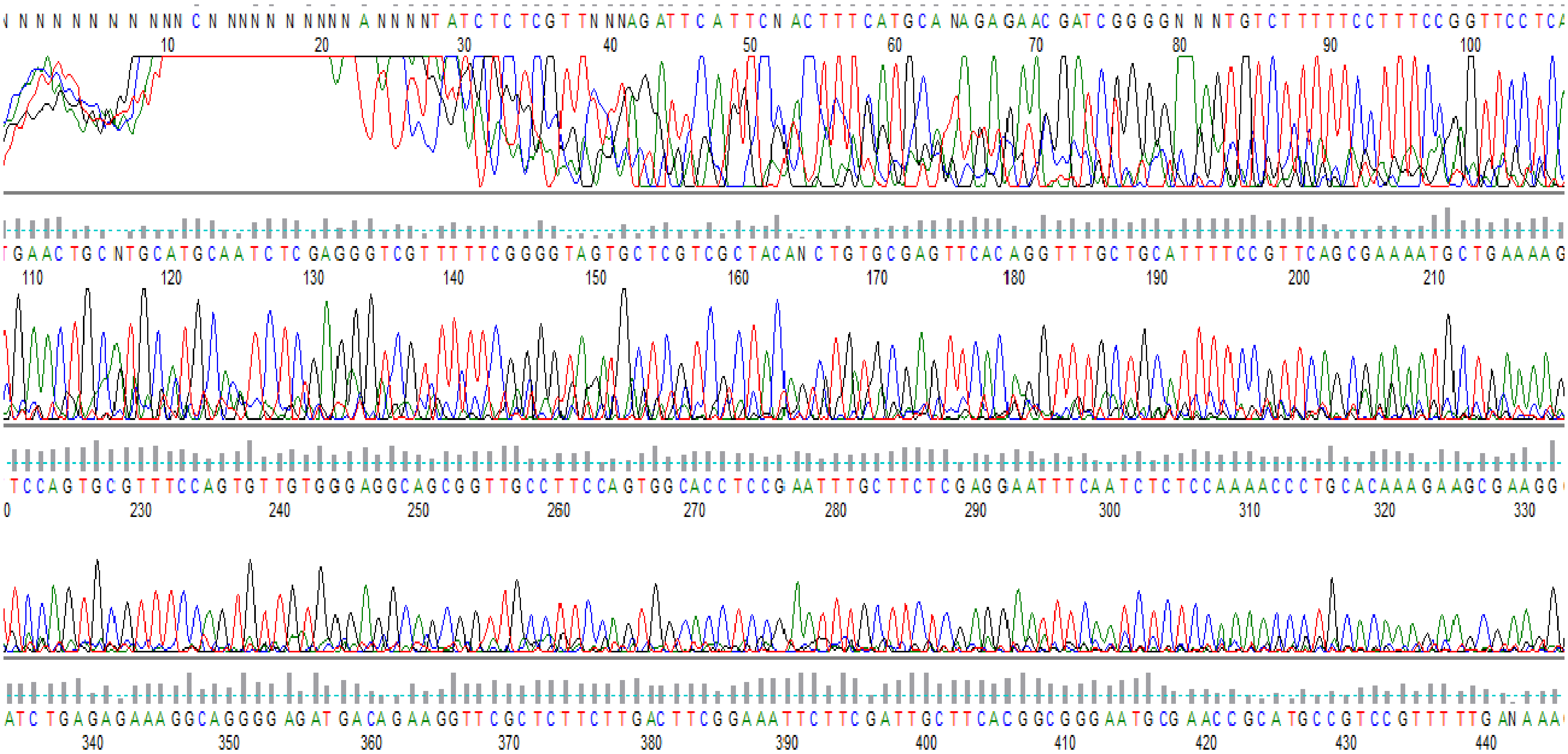
Sequence analysis of the Apico marker (subject 829) was identified by BLAST as *T. gondii* TgCatAu_9 strain (91%).



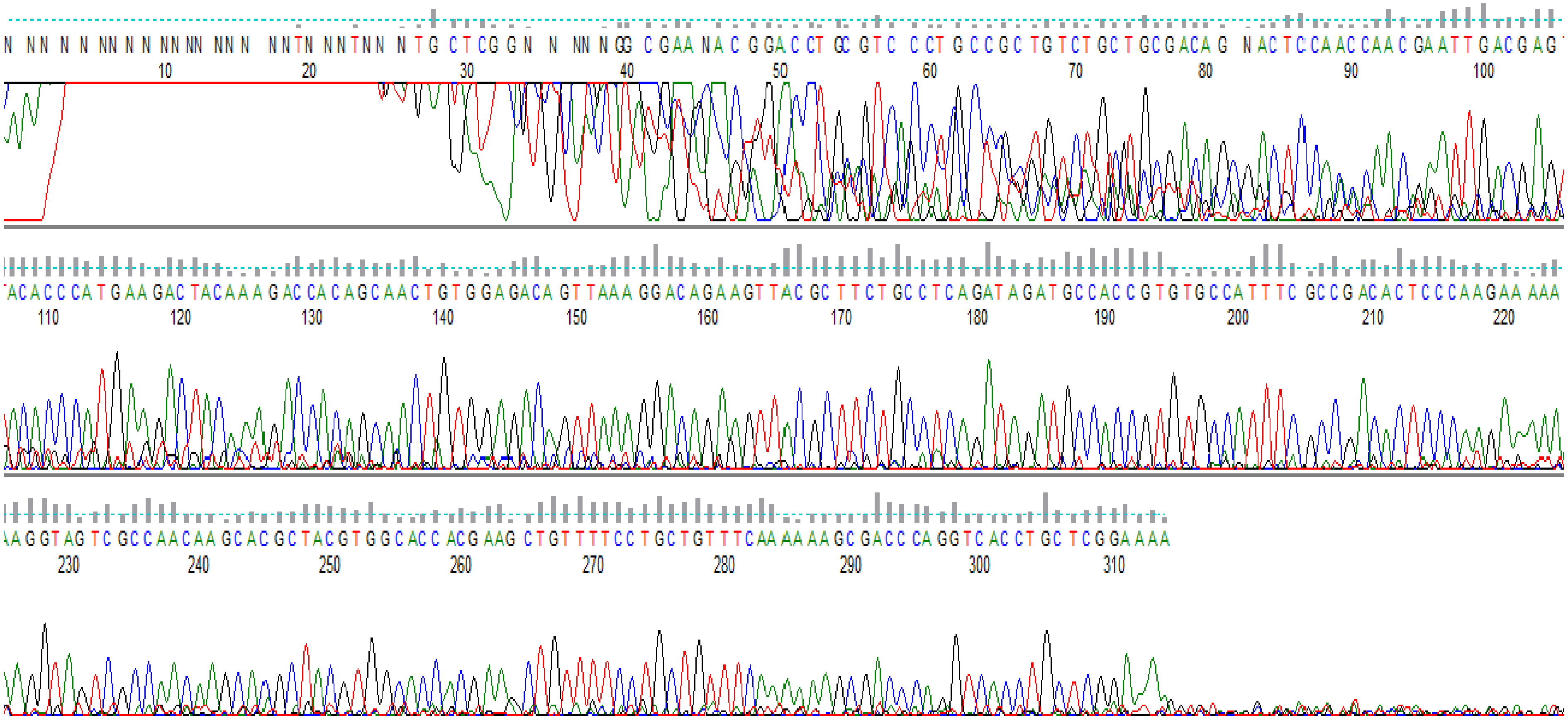
Sequence analysis of the BTUB marker (subject 1067) was identified by BLAST as *T. gondii* strain DEG or BEVERLEY (94%).



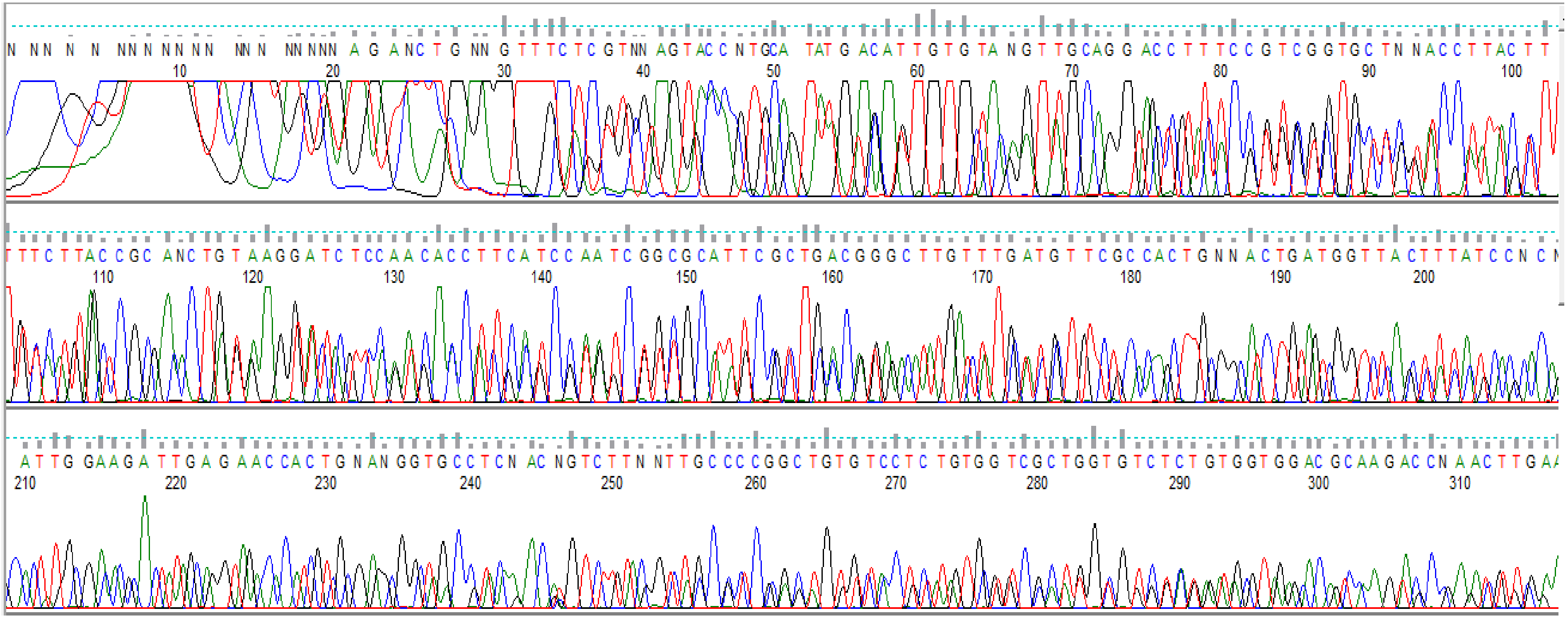
Sequence analysis of the C22-8 marker (subject 1032) was identified by BLAST as *T. gondii* strain VEG or TgCatAu_7 (96%).



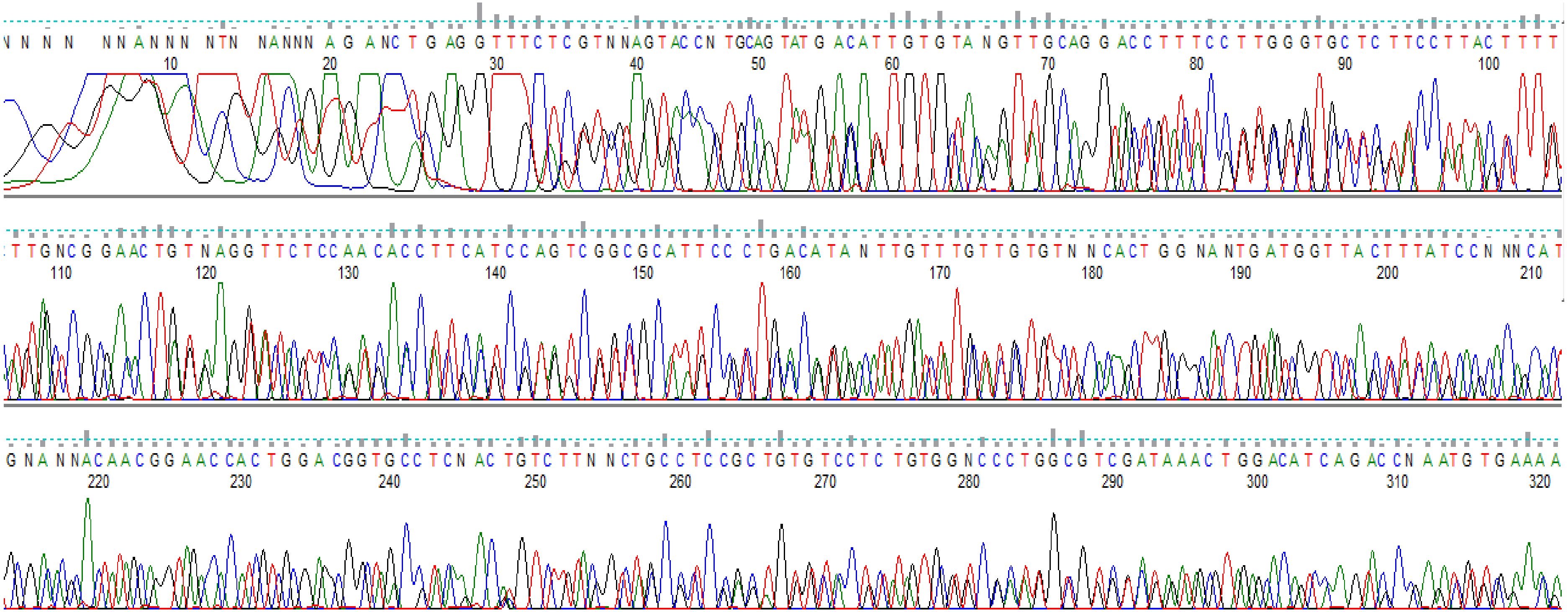
Sequence analysis of the GRA6 marker (subject 674) was identified by BLAST as *T. gondii* AF07/02U strain (99%).



Sequence analysis of the SAG2 3/ marker (subject 918) was identified by BLAST as *T. gondii* Me-49 strain



Sequence analysis of the SAG2 5/ marker (subject 918) was identified by BLAST as *T. gondii* Me-49 strain



List of publications

Bajnok, J.; Boyce, K.; Rogan, M.Q T.; Craig, P. S.; Lun, Z. R. Hide, G. (2014). Prevalence of *Toxoplasma gondii* in localized populations of *Apodemus sylvaticus* is linked to population genotype not to population location. Parasitology. 142 (5): 680-690

Morger J, Bajnok J, Boyce K, Craig PS, Rogan MT, Lun ZR, Hide G, Tschirren B. (2014). Naturally occurring Toll-like receptor 11 (TLR11) and Toll-like receptor 12 (TLR12) polymorphisms are not associated with *Toxoplasma gondii* infection in wild wood mice. Infection, Genetics and Evolution. 26:180-4.